

ANNALES
HENRI LEBESGUE

ARSENIY AKOPYAN

VLADISLAV VYSOTSKY

LARGE DEVIATIONS OF CONVEX
HULLS OF PLANAR RANDOM
WALKS AND BROWNIAN MOTIONS

GRANDES DÉVIATIONS POUR
L'ENVELOPPE CONVEXE DE MARCHES
ALÉATOIRES ET DE MOUVEMENTS
BROWNIENS DANS LE PLAN

Keywords: Random walk, Brownian motion, Wiener process, Lévy process, convex hull, large deviations, perimeter, area, mean width, rate function, non-convex rate function, radial minimum, radial maximum, Legendre–Fenchel transform, convex conjugate.

2020 *Mathematics Subject Classification:* 60D05, 60F10, 60G50, 26B25, 52A22, 60G70.

DOI: <https://doi.org/10.5802/ahl.100>

(*) This paper was written when AA was affiliated to IST Austria. His work was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n°[291734] and European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 78818 Alpha). This paper was partially written when VV was affiliated to Imperial College London, where his work was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n°[628803].

ABSTRACT. — We prove large deviations principles (LDPs) for the perimeter and the area of the convex hull of a planar random walk with finite Laplace transform of its increments.

We give explicit upper and lower bounds for the rate function of the perimeter in terms of the rate function of the increments. These bounds coincide and thus give the rate function for a wide class of distributions which includes the Gaussians and the rotationally invariant ones. For random walks with such increments, large deviations of the perimeter are attained by the trajectories that asymptotically align into line segments. However, line segments may not be optimal in general.

Furthermore, we find explicitly the rate function of the area of the convex hull for random walks with rotationally invariant distribution of increments. For such walks, which necessarily have zero mean, large deviations of the area are attained by the trajectories that asymptotically align into half-circles. For random walks with non-zero mean increments, we find the rate function of the area for Gaussian walks with drift. Here the optimal limit shapes are elliptic arcs if the covariance matrix of increments is non-degenerate and parabolic arcs if otherwise.

The above results on convex hulls of Gaussian random walks remain valid for convex hulls of planar Brownian motions of all possible parameters. Moreover, we extend the LDPs for the perimeter and the area of convex hulls to general Lévy processes with finite Laplace transform.

RÉSUMÉ. — Nous montrons des principes de grandes déviations pour le périmètre et l'aire de l'enveloppe convexe d'une marche aléatoire planaire dont les incréments ont une transformée de Laplace finie.

Nous donnons des bornes inférieures et supérieures explicites pour la fonction de taux pour le périmètre, en termes de la fonction de taux pour les incréments. Pour une large classe de distributions incluant les distributions gaussiennes et les distributions invariantes par rotation, ces bornes coïncident et donnent donc la fonction de taux exacte. Pour des marches aléatoires avec de tels incréments, les grandes déviations pour le périmètre sont atteintes par les trajectoires qui se comportent asymptotiquement comme des segments de droite. Cependant, les segments de droite ne sont pas optimaux en général.

De plus, nous trouvons explicitement la fonction de taux pour l'aire de l'enveloppe convexe des marches aléatoires à incréments invariants par rotation. Pour de telles marches, qui sont nécessairement centrées, les grandes déviations pour l'aire sont réalisées par les trajectoires qui se comportent asymptotiquement comme des demi-cercles. Pour des marches aléatoires avec des incréments non centrés, nous donnons la fonction de taux pour l'aire pour des marches gaussiennes non centrées. En ce cas, les trajectoires optimales sont des arcs elliptiques si la matrice de covariance des incréments est non dégénérée, et des arcs paraboliques sinon.

Les résultats précédents sur les enveloppes convexes de marches aléatoires gaussiennes restent valides pour les enveloppes convexes de mouvements browniens plans pour tous les paramètres. De plus, nous étendons le principe de grande déviation pour le périmètre et l'aire des enveloppes convexes à des processus de Lévy généraux avec transformée de Laplace finie.

1. Introduction

Let $(S_k)_{k \geq 1}$, where $S_k = X_1 + \dots + X_k$, be a planar random walk with independent identically distributed increments X_1, X_2, \dots . We assume that the expectation of X_1 exists and is finite, and put $\mu := \mathbb{E}X_1$. We are interested in the perimeter P_n and the area A_n of the convex hull $C_n := \text{conv}(0, S_1, \dots, S_n)$ of the first n steps of the random walk, including the origin. Here, by definition, the perimeter of a line segment is its doubled length.

All of our results remain valid for the convex hulls $\text{conv}(S_1, \dots, S_n)$ but it is more natural to consider hulls of the form C_n , which allow remarkably simple formulas for

their expected perimeters and areas. In fact, Spitzer and Widom [SW61] proved⁽¹⁾ that

$$(1.1) \quad \mathbb{E}P_n = 2 \sum_{k=1}^n \frac{\mathbb{E}|S_k|}{k},$$

where by $|\cdot|$ we denote the Euclidean norm. This implies that $\mathbb{E}P_n/n \rightarrow 2|\mu|$ as $n \rightarrow \infty$, by the law of large numbers and uniform integrability of $(S_k/k)_{k \geq 1}$. Moreover, $P_n/n \rightarrow 2|\mu|$ a.s. by McRedmond and Wade [MW18].

Wade and Xu [WX15a] showed (developing the ideas introduced by Snyder and Steele [SS93]) that if $\mu \neq 0$ and $\mathbb{E}|X_1|^2 < \infty$, then $\text{Var}(P_n)/n \rightarrow 4\sigma_\mu^2/|\mu|^2$, where $\sigma_\mu^2 := \mathbb{E}(\mu \cdot (X_1 - \mu))^2$ and ‘ \cdot ’ denotes the scalar product. Here $\sigma_\mu^2 > 0$ unless the trajectory of $(S_k)_{k \geq 1}$ is the graph of a zero-mean one-dimensional random walk. With the exception of this degenerate case, the variance of the perimeter grows linearly (when $\mathbb{E}|X_1|^2 < \infty$), and moreover, the sequence $(P_n)_{n \geq 1}$ satisfies a central limit theorem for $\mu \neq 0$ (see [WX15a]) and a limit theorem under the scaling $n^{-1/2}$ for $\mu = 0$ (see Wade and Xu [WX15b]). The latter result follows naturally from the invariance principle using the continuous mapping theorem. The degenerate case $\mu \cdot (X_1 - \mu) = 0$ a.s. is more tricky. Alsmeyer et al. [AKMV20] proved that $\text{Var}(P_n) = O(\log n)$ under $\mathbb{E}|X_1|^3 < \infty$; it is likely that $\text{Var}(P_n)$ may grow super-logarithmically (contradicting the corresponding conjecture in [WX15a]) when $\mathbb{E}|X_1|^3 = \infty$ and $\mathbb{E}|X_1|^2 < \infty$, but a proof is still missing. Yet there is no central limit theorem for $(P_n)_{n \geq 1}$, although Alsmeyer et al. [AKMV20] established the ones for the lengths of the convex minorant and the concave majorant of n -step one-dimensional random walks (the sum of these quantities is P_n ; their variances may grow polynomially when $\mathbb{E}|X_1|^3 = \infty$ but not much is known about their correlation).

The Spitzer–Widom formula (1.1) admits various generalizations to higher dimensions, including explicit formulas for the expected mean width, surface area, volume, and other intrinsic volumes of the convex hulls, see Barndorff-Nielsen and Baxter [BNB63] or Vysotsky and Zaporozhets [VZ18]. In particular, for the area of the convex hull of a planar random walk,

$$(1.2) \quad \mathbb{E}A_n = \frac{1}{2} \sum_{\substack{j, k \geq 1 \\ j+k \leq n}} \frac{\mathbb{E}|\det[S_j, S'_k]|}{jk},$$

where $(S'_k)_{k \geq 1}$ is an independent copy of $(S_k)_{k \geq 1}$. Furthermore, the invariance principle naturally implies (see Wade and Xu [WX15b]) that if $\mathbb{E}|X_1|^2 < \infty$, then the sequence $(A_n)_{n \geq 1}$ satisfies a limit theorem under the scaling $n^{-3/2}$ for $\mu \neq 0$ and n^{-1} for $\mu = 0$.

In this paper we study large deviations probabilities for the perimeter and the area of the convex hull of the random walk. This describes very atypical behaviour of these quantities, as opposed to the results above on their typical behaviour. In

⁽¹⁾[SW61] proved formula (1.1) under the assumption $\mathbb{P}(u \cdot S_k = 0) = 0$ for every $k \in \mathbb{N}$ and non-zero $u \in \mathbb{R}^2$, which can be dropped using a simple approximation argument based on the fact that the perimeter is a continuous functional on the space of compact convex sets equipped with the Hausdorff distance. This observation also applies to equality (1.2).

particular, we will consider the logarithmic asymptotics of $\mathbb{P}(P_n \geq 2xn)$ for $x > |\mu|$ and $\mathbb{P}(P_n \leq 2xn)$ for $x < |\mu|$, and $\mathbb{P}(A_n \geq an^2)$ for $a > 0$. We will also describe the limit shape of the trajectories, scaled by the factor of n^{-1} in both time and space (this explains the scalings of P_n and A_n), resulting in such large deviations probabilities.

To the best of our knowledge, there is only one rigorous result in this direction. Snyder and Steele [SS93] obtained the following non-sharp concentration inequality for the perimeter for random walks with bounded increments: if $|X_1| \leq M$ a.s. for some $M > 0$, then

$$\mathbb{P}(|P_n - \mathbb{E}P_n| \geq xn) \leq 2 \exp\left(-x^2n / (8\pi^2M^2)\right), \quad x \geq 0.$$

Claussen et al. [CHM15] gave a numerical analysis of atypically large values of the perimeter and the area of the convex hull and concluded that these quantities “seem... to obey a large deviations principle” (*LDP*, in short) for random walks with standard Gaussian increments. There are few follow-up numerical papers on related questions by the same group of authors.

The other results on atypical behaviour of convex hulls include the works by Khoshnevisan [Kho92] and Kuelbs and Ledoux [KL98], who considered a.s. superior limits of monotone functionals of convex hulls (including the perimeter and the area) of zero-mean finite-variance random walks and standard Brownian motions scaled as in the law of iterated logarithm.

The simulation-based conclusions of [CHM15] in fact easily follow (see Section 4.2 below) from the contraction principle applied to Mogulskii’s LDP for trajectories of random walks with finite Laplace transform of their increments. The main task is to obtain *explicitly* the rate functions in these LDP’s in terms of the rate function of the increments. For the perimeter, we found the rate function (Corollary 2.10) for a wide class of random walks (see Proposition 2.9) including all Gaussian walks (this is not at all an expected result), and also gave the upper and the lower bound valid for general walks with finite Laplace transform of increments (Theorem 2.3). For the area, we found the rate function for random walks that have rotationally invariant distributions of increments with finite Laplace transform (Theorem 2.11) and for Gaussian random walks with arbitrary drift (Theorem 2.13 and Proposition 2.15). In all these results we identified the asymptotic form of optimal trajectories of the walk resulting in the large deviations.

Furthermore, we extended the above results on random walks, which have increments in discrete time, to convex hulls of planar Lévy processes (Theorem 2.16) with finite Laplace transform, including Brownian motions. Convex hulls of general Lévy processes were studied e.g. by Molchanov and Wespi [MW16].

Lastly, we extended the LDPs for P_n and A_n to random walks whose increments have Laplace transform finite only in a neighbourhood of zero (Proposition 4.1).

The paper is organized as follows. In Section 2.1 we introduce notation and in particular, define the *radial minimum rate function* and state its properties. In Sections 2.2 and 2.3 we present our main results on large deviations for the perimeter and the area of the convex hull of a planar random walk. The continuous-time

counterparts are given in Section 2.4. Further generalizations are discussed in Section 2.5. In Section 3 we prove basic properties of the radial minimum rate function for general increments of the walk and its convexity for Gaussian walks. Section 4.1 contains essentials on large deviations relevant to this paper. The proofs of our LDPs for the perimeter and the area (for random walks), including computations of the rate functions, are given in Section 4.2. The core parts of these computations are unified by the use of geometric inequalities of isoperimetric type. The proofs for Lévy processes are in Section 4.3. Finally, in Section 4.4 we give a partial result for walks with Laplace transform of increments finite only in a neighbourhood of zero.

2. Main results

2.1. Notation

Recall that the *Legendre–Fenchel transform* or the *convex conjugate* of a function $F : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ (where $d \geq 1$) with a non-empty *effective domain* $\mathcal{D}_F := \{u : F(u) < \infty\}$ is the function $F^* : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$F^*(v) := \sup_{u \in \mathbb{R}^d} (u \cdot v - F(u)), \quad v \in \mathbb{R}^d.$$

The conjugate function F^* is convex and lower semi-continuous on \mathbb{R}^d ; F itself does not need to be convex. Recall that any convex function F is continuous on the relative interior $\text{rint } \mathcal{D}_F$ of its effective domain (Rockafellar [Roc70, Theorem 10.1]) so the property of lower semi-continuity is needed to characterize F only near the relative boundary of \mathcal{D}_F . By $\text{conv } F$ we denote the *largest convex minorant* or the *convex hull* of F , i.e. the convex function with the epigraph $\text{conv}(\text{epi } F)$, which is a subset of \mathbb{R}^{d+1} . Thus, we use the notation “conv” both for functions and sets.

The *cumulant generating function* $K(u) := \log \mathbb{E}e^{u \cdot X_1}$ is convex by Jensen’s inequality and satisfies $K(0) = 0$. Its convex conjugate $I := K^*$ is the *rate function* of X_1 . This function satisfies $I(\mu) = 0$ and is non-negative, lower semi-continuous, and continuous on $\text{rint } \mathcal{D}_I$, where rint stands for the *relative interior* (taken in the induced topology of the affine hull of \mathcal{D}_I). In the main results of this paper (namely, the LDPs for P_n and A_n) we assume that the Laplace transform of the increments $\mathcal{L}(u) := \mathbb{E}e^{u \cdot X_1}$ is finite for all $u \in \mathbb{R}^2$. For example, this is trivially true when the support of X_1 is bounded. Under this assumption, K is infinitely differentiable on \mathbb{R}^2 and I is *strictly* convex on its effective domain \mathcal{D}_I ; see Barndorff-Nielsen [BN78, Corollary 7.1] and Vysotsky [Vys21b, Corollary 2.12].

The effective domain of I is known to satisfy (see [Vys21b, Proposition 1.1.a])

$$(2.1) \quad \text{rint} \left(\text{conv}(\text{supp}(X_1)) \right) \subset \mathcal{D}_I \subset \text{cl} \left(\text{conv}(\text{supp}(X_1)) \right),$$

where $\text{supp}(X_1)$ is the topological support of the distribution of X_1 . Furthermore, put

$$r_{\min} := \inf \{|u| : u \in \text{conv}(\text{supp}(X_1))\}, \quad r_{\max} := \sup \{|u| : u \in \text{supp}(X_1)\}.$$

Note that $r_{\min} \leq |\mu| \leq r_{\max}$, where the second inequality is strict unless $X_1 = \mu$ a.s. and the first inequality is strict unless $\mu \cdot (X_1 - \mu) = 0$ a.s.

On occasions, we will give general statements assuming that the random vector X_1 takes values in \mathbb{R}^d with an arbitrary $d \geq 1$ rather than merely in \mathbb{R}^2 . With no risk of confusion, in such cases in $\mathcal{L}(u) = \mathbb{E}e^{u \cdot X_1}$ we take $u \in \mathbb{R}^d$ and understand I, K , etc. accordingly. Then we will usually assume that X_1 satisfies merely the *Cramér moment assumption* $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$.

Define the *radial maximum* and *radial minimum* functions

$$(2.2) \quad \underline{I}(r) := \inf_{\ell \in \mathbb{S}^{d-1}} I(r\ell), \quad \bar{K}(p) := \sup_{\ell \in \mathbb{S}^{d-1}} K(p\ell), \quad p, r \geq 0,$$

where \mathbb{S}^{d-1} stands for the unit sphere in \mathbb{R}^d centred at 0, and put $\underline{I}(r) := \infty$ and $\bar{K}(p) := \infty$ for $p, r < 0$. Note that the function \underline{I} admits the following geometric interpretation: the epigraph of $\underline{I}(|v|)$ is the union of all rotations of the epigraph of $I(v)$ about the vertical axis. Clearly, the supremum and the infimum above are always attained at some points since the Laplace transform is continuous, I is lower semi-continuous, and spheres in \mathbb{R}^d are compact. Thus, the respective sets of minimal and maximal directions

$$(2.3) \quad \underline{\Lambda}_r := \operatorname{argmin}_{\ell \in \mathbb{S}^{d-1}} I(r\ell), \quad \bar{\Lambda}_p := \operatorname{argmax}_{\ell \in \mathbb{S}^{d-1}} K(p\ell), \quad r, p \geq 0$$

are always non-empty. Note that the argmax will not change if we replace K by \mathcal{L} in (2.3).

Recall that a point u in a convex set $C \subset \mathbb{R}^d$ is called *extreme* if there is no way to express $u = \alpha u_1 + (1 - \alpha)u_2$ for some $u_1, u_2 \in C$ and $\alpha \in (0, 1)$ except by taking $u_1 = u_2 = u$. Every extreme point of C belongs to the *relative boundary* $\partial_{\text{rel}}C$ of C , defined by $\partial_{\text{rel}}C := C \setminus \text{rint } C$. An extreme point u of a convex set $C \subset \mathbb{R}^d$ is called *exposed* if $C \cap L = \{u\}$ for some hyperplane L supporting C .

The radial minimum rate function \underline{I} and the sets of minimal directions $\underline{\Lambda}_r$ appear in most of our results on the perimeter of the convex hull. Let us state some of their properties. Let us agree that by $[|\mu|, r_{\max}]$ we will mean the half-line $[|\mu|, \infty)$ if $r_{\max} = \infty$.

LEMMA 2.1. — Assume that X_1 is a random vector in \mathbb{R}^d , where $d \geq 1$, such that $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$.

- (a) The effective domain $\mathcal{D}_{\underline{I}}$ of \underline{I} is an interval that satisfies $\text{int } \mathcal{D}_{\underline{I}} = (r_{\min}, r_{\max})$;
- (b) The function \underline{I} is lower semi-continuous; satisfies $\underline{I}(|\mu|) = 0$; is strictly decreasing and convex (also strictly if $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^d$) on $[r_{\min}, |\mu|]$; and is strictly increasing on $[|\mu|, r_{\max}]$;
- (c) Suppose that \underline{I} is discontinuous at a point $x \in [|\mu|, r_{\max}]$. Then for any $\ell \in \underline{\Lambda}_x$, $x\ell$ is an exposed point of $\mathcal{D}_{\underline{I}}$ and $\underline{I}(x) = -\log \mathbb{P}(X_1 = x\ell) < \infty$.
- (d) For any $r \in (r_{\min}, |\mu|]$, the set $\underline{\Lambda}_r$ contains a unique element, which we denote by ℓ_r .

Combining Part (c) with the second inclusion in (2.1) gives:

COROLLARY 2.2. — \underline{I} is continuous on (r_{\min}, ∞) if $\mathbb{P}(X_1 = u) = 0$ for any $u \in \partial_{\text{rel}}(\text{conv}(\text{supp}(X_1)))$.

We stress that the function \underline{I} may be discontinuous on its effective domain, and may be non-convex even if it is continuous; see Remark 2.4 in Section 2.2 and Example 3.1 in Section 3.2.

2.2. Large deviations of the perimeter

Denote by $AC_0([0, 1]; \mathbb{R}^2)$ the set of coordinate-wise absolutely continuous functions h on $[0, 1]$ such that $h(0) = 0$. We will occasionally refer to functions from $[0, 1]$ to \mathbb{R}^2 as (planar) *curves* or *trajectories*. Denote by $\text{im}(\cdot)$ the image of a function, that is the set of its values as the argument varies over the effective domain. Let $P(C)$ denote the perimeter of a non-empty convex set $C \subset \mathbb{R}^2$, so $P_n = P(C_n)$.

We now state our first main result.

THEOREM 2.3. — *Assume that X_1 is a random vector in \mathbb{R}^2 such that $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$.*

(1) *The sequence $(P_n/(2n))_{n \geq 1}$ satisfies the LDP in \mathbb{R} with speed n and the tight rate function*

$$(2.4) \quad \mathcal{J}_P(x) := \min_{\substack{h \in AC_0([0, 1]; \mathbb{R}^2) \\ P(\text{conv}(\text{im } h)) = 2x}} \int_0^1 I(h'(t)) dt.$$

This function shares the properties of \underline{I} stated in Parts (a) and (b) of Lemma 2.1.

(2) *We have $\mathcal{J}_P = \underline{I}$ on $[0, |\mu|]$ and $\text{conv } \underline{I} \leq \mathcal{J}_P \leq \underline{I}$ on $[|\mu|, \infty)$, and a bit more:*

$$(2.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(P_n \leq 2xn) = \begin{cases} -\underline{I}(x), & x \in (r_{min}, |\mu|] \\ \log \mathbb{P}(|X_1| = r_{min}), & x = r_{min} \end{cases}$$

and for any $x \in [|\mu|, r_{max}]$,

$$(2.6) \quad \begin{aligned} -\underline{I}(x) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(P_n \geq 2xn) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(P_n \geq 2xn) \leq -\text{conv } \underline{I}(x). \end{aligned}$$

(3) *For any $\varepsilon > 0$, we have*

$$(2.7) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{0 \leq k \leq n} \left| \frac{S_k}{n} - \frac{k}{n} x \ell_x \right| \leq \varepsilon \mid P_n \leq 2xn \right) = 1, \quad x \in (r_{min}, |\mu|],$$

where ℓ_x was defined in Lemma 2.1(d), and

$$(2.8) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{0 \leq k \leq n} \left| \frac{S_k}{n} - h(k/n) \right| \leq \varepsilon \text{ for some } h \in H_P(x) \mid P_n \geq 2xn \right) = 1, \\ x \in [|\mu|, r_{max}),$$

where $H_P(x)$ denotes the set of minimizers in (2.4). If $\underline{I}(x) = \text{conv } \underline{I}(x)$ for an $x \in (r_{min}, r_{max})$, then $H_P(x) = \{t \mapsto tx\ell\}_{\ell \in \Delta_x}$.

We refer to the elements of the sets $H_P(x)$ as the *optimal trajectories* (for the perimeter).

Let us give a few comments. The limit shape results (2.7) and (2.8) mean that if $\underline{I}(x) = \text{conv } \underline{I}(x)$, then large deviations of the perimeter are attained on trajectories that asymptotically align into line segments and move with constant speed. Note that under $\underline{I}(x) = \text{conv } \underline{I}(x)$, equality (2.8) does not assert that every direction in $\underline{\Lambda}_x$ can be attained. This equality means that the epigraph of \underline{I} admits a support line at the point $(x, \underline{I}(x))$. This is true for every x iff \underline{I} is convex, in which case the rate function in the LDP for the perimeter is $\mathcal{J}_P = \underline{I}$. In Proposition 2.6 below we will provide a tractable condition, stated directly in terms of the Laplace transform of increments, for checking the equality $\underline{I}(x) = \text{conv } \underline{I}(x)$ for a given x . Moreover, this proposition relates the sets of optimal directions $\underline{\Lambda}_x$ in (2.8) to more tractable sets $\bar{\Lambda}_p$ defined in terms of the Laplace transform; cf. (2.3).

The idea of our proof of Theorem 2.3 is as follows. Equalities (2.4), (2.7), and (2.8) follow from an LDP for trajectories of random walks combined with the contraction principle. If $\underline{I}(x) = \text{conv } \underline{I}(x)$, an additional geometric argument yields that the set $H_P(x)$ consists of curves of minimal length with the fixed perimeter $2x$ of their convex hull. A known geometric result (Corollary A.2 in the Appendix) asserts that the image of such a curve is a line segment of length x . This curve must move with constant speed by strict convexity of I .

With this geometric optimality property of line segments, it is tempting to assume that $\mathcal{J}_P = \underline{I}$ and we always have $H_P(x) = \{t \mapsto tx\ell\}_{\ell \in \underline{\Lambda}_x}$ in (2.8). However, in general,

the optimal trajectories are not necessarily linear.

Hence it may be that $\mathcal{J}_P \neq \underline{I}$, as shown in Example 2.5, which follows the next remark.

Remark 2.4. — For $\mu \neq 0$, the upper bound in Part (2) of Theorem 2.3 can be improved to $\mathcal{J}_P \leq \underline{I}_0$ on $[|\mu|, \infty)$, where $I_0 := \text{conv}(\mathbb{1}_{\mathbb{R} \setminus \{0\}} \cdot I)$ is the function with the epigraph $\text{conv}(\text{epi } I \cup \{0\})$. This follows from considering the set of trajectories

$$D := \left\{ h \in AC_0[0, 1] : h' = v \text{ on } [0, s], h' = \mu \text{ on } [s, 1] \text{ for some } s \in (0, 1), v \in T \right\},$$

where $T := \{v \in \mathcal{D}_I : I(tv) \geq tI(v) \text{ for any } t \in \mathbb{R}\}$, whose energies satisfy

$$\int_0^1 I(h') dt = sI(v) + (1 - s)I(\mu) = sI(v) < I(sv).$$

Note in passing that it is easy to check that $T = (K \circ \nabla I)^{-1}(0)$ if ∇I is defined on \mathbb{R}^2 .

We actually have $\mathcal{J}_P = \underline{I}_0$ on $[|\mu|, \infty)$ if the distribution of X_1 is supported on the straight line $\mu\mathbb{R}$ (and satisfies $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$). Even in this degenerate case \underline{I}_0 differs from \underline{I} on $[|\mu|, \infty)$ if there is an $a > 1$ such that $K(-a\mu) = 0$ and $I(a\mu) > I(-a\mu)$ (cf. Example 2.5). There are only two directions, therefore both optimal trajectories, one of which belongs to D , start *moving backwards* at some moment. This is very counter-intuitive! In this case \mathcal{J}_P is non-convex, and so is \underline{I} . Unfortunately, the case $\mathbb{P}(X_1 \in \mu\mathbb{R}) = 1$ is the only type of distribution of the increments with a possibly non-convex \mathcal{J}_P where we found \mathcal{J}_P explicitly.

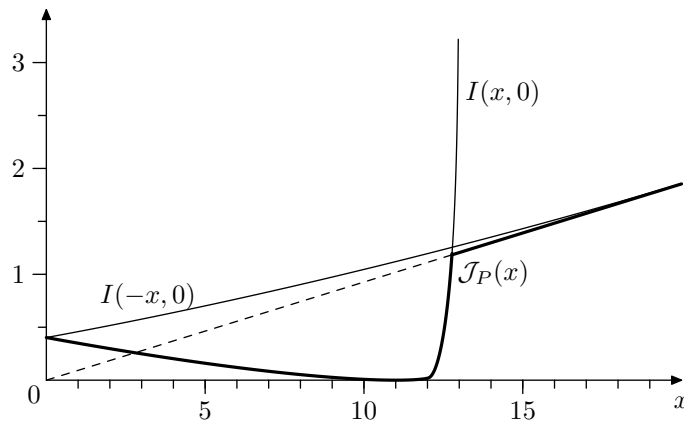


Figure 2.1. $\underline{I}_0(x) < \underline{I}(x)$ for $x \in [12.77, 21.88]$.

Example 2.5 (Non-linear optimal trajectories). — Consider the distribution of X_1 given by the mixture of the uniform distribution on the line segment $[11, 13] \times \{0\}$ taken with weight $49/50$ and the delta distribution at $(-38, 0)$ taken with weight $1/50$; see Figure 2.4.

We now turn our attention to the function $\text{conv } \underline{I}$. The following result gives a simple description of $\text{conv } \underline{I}$ directly in terms of \bar{K} and a condition when it equals \underline{I} at a given point. Denote by $(\cdot)'_+$ and $(\cdot)'_-$ respectively right and left derivatives of a function of real argument.

PROPOSITION 2.6. — Assume that X_1 is a random vector in \mathbb{R}^d , $d \geq 1$, such that $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^d$. Then

(a) \bar{K} is an increasing convex function on $[0, \infty)$ satisfying $\bar{K}'_+(0) = |\mu|$ and

$$\text{conv } \underline{I} = (\bar{K})^* \text{ on } [|\mu|, \infty);$$

(b) If $r \in \text{cl}(\text{im}(\bar{K}'))$ (and $r \geq |\mu|$), then $\underline{I}(r) = \text{conv } \underline{I}(r) < \infty$;

(c) For any $p \in (0, \infty)$, the one-sided derivatives satisfy

$$\bar{K}'_+(p) = \max_{\ell \in \Lambda_p} |\nabla K(p\ell)| \quad \text{and} \quad \bar{K}'_-(p) = \min_{\ell \in \Lambda_p} |\nabla K(p\ell)|.$$

(d) If $p \in (0, \infty)$ and $r \geq |\mu|$ are such that $\bar{K}'(p) = r$, then $\underline{\Lambda}_r = \bar{\Lambda}_p$.

The main result of the proposition is Part (a), which relates the radial maximum of the logarithmic Laplace transform K to the radial minimum of its convex conjugate I . This assertion is actually a general fact valid for arbitrary convex functions; see Proposition 3.2 of Section 3.3, which yields a stronger version of Proposition 2.6 under the Cramér moment assumption. Part (b) is an easy consequence of Part (a) and the well-known fact that the Legendre–Fenchel transform maps kinks of a convex function (in our case, \bar{K}) into linear segments of its convex conjugate. Part (c), which clarifies the possible reason of non-differentiability of \bar{K} , follows by a standard application of the method of Lagrange multipliers. Part (d), which follows naturally from Parts (a)-(c), claims that the slowest directions of I are exactly the fastest directions of K (equivalently, of \mathcal{L}) at the corresponding radii.

The main use of Proposition 2.6 is through its following corollaries.

COROLLARY 2.7. — \underline{I} is strictly convex on $\mathcal{D}_{\underline{I}}$ if \bar{K} is differentiable on $(0, \infty)$.

COROLLARY 2.8. — \bar{K} is differentiable if there exists a continuous mapping $\ell : (0, \infty) \rightarrow \mathbb{S}^{d-1}$ such that $\ell(p) \in \bar{\Lambda}_p$ for any $p \in (0, \infty)$.

These claims follow easily from Parts (b) and (c), respectively, using convexity of \bar{K} ; see Section 3.3. We do not assume that $\ell(p)$ is differentiable, otherwise Corollary 2.8 becomes trivial.

We now present a few types of distributions with convex radial minimum rate function \underline{I} . By $\|\cdot\|$ we denote the largest eigenvalue of a symmetric real matrix.

PROPOSITION 2.9. — Assume that X_1 is a random vector in \mathbb{R}^d , $d \geq 1$. Then the function \underline{I} is convex in either of the following cases:

- (a) $X_1 = AY_1 + \mu$, where Y_1 is a random vector in \mathbb{R}^k , $1 \leq k \leq d$, with rotationally invariant distribution and A is a $d \times k$ real matrix such that $AA^\top \mu = \|AA^\top\|\mu$;
- (b) X_1 is Gaussian.

COROLLARY 2.10. — For the above types of distributions, we have $\mathcal{J}_P = \underline{I}$ (cf. Theorem 2.3).

We regard that Corollary 2.10 gives the rate function \mathcal{J}_P explicitly, since finding the radial minimum \underline{I} is a standard optimization problem (which is much simpler than (2.4)) solvable using the method of Lagrange multipliers.

We now comment on Proposition 2.9. Note that in Case (a), the matrix AA^\top is proportional to the covariance matrix of X_1 given by $\Sigma := \mathbb{E}(X_1 X_1^\top) - \mu \mu^\top$. The assumption on A is always satisfied if $\mu = 0$.

For Case (a), convexity of \underline{I} follows rather directly from that of I . Our proof for Case (b) rests on Corollaries 2.7 and 2.8 and uses properties of quadric curves to construct a path $\ell(p)$.

The condition in Corollary 2.8 is trivially satisfied if the set $\cap_{p>0} \bar{\Lambda}_p$ is non-empty, i.e. there exists a direction that maximizes the Laplace transform at all radii. This rather restrictive assumption naturally holds true for either linearly transformed or shifted rotationally invariant distributions of increments. Both cases are covered by Case (a) of Proposition 2.9, where $\cap_{p>0} \bar{\Lambda}_p$ is the set of maximal eigenvectors of Σ of unit length if $\mu = 0$ and $\cap_{p>0} \bar{\Lambda}_p = \{\mu/|\mu|\}$ if $\mu \neq 0$. For general affine transforms (i.e. compositions of linear transforms and translations) of rotationally invariant distributions, we were able to prove convexity of \underline{I} only for Gaussian distributions, as per Case (b). We will see that here $\cap_{p>0} \bar{\Lambda}_p$ is empty unless the Gaussian distribution of X_1 is degenerate or X_1 satisfies the assumptions of Case (a).

2.3. Large deviations of the area

The main geometric argument used in our proof of Theorem 2.3 to find the rate function in the LDP for the perimeter can be applied directly to obtain the rate function in the LDP for the area A_n of the convex hull. The problem reduces to

finding a planar curve of the unit length that maximizes the area of its convex hull. This question is known as one of Ulam’s problems. Although it is very similar to the classical Dido problem and of course has the same answer that the curve is a half-circle (Moran [Mor46]), it appears that this Ulam problem does not allow an easy solution by reduction. The corresponding isoperimetric inequality easily yields the following result.

Denote by $A(C)$ the area of a non-empty convex set $C \subset \mathbb{R}^2$, so $A_n = A(C_n)$.

THEOREM 2.11. — Assume that X_1 is a random vector in \mathbb{R}^2 such that $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$.

(1) The sequence $(A_n/n^2)_{n \geq 1}$ satisfies an LDP with speed n and the tight rate function

$$(2.9) \quad \mathcal{J}_A(a) := \min_{\substack{h \in AC_0([0,1]; \mathbb{R}^2) \\ A(\text{conv}(\text{im } h)) = a}} \int_0^1 I(h'(t)) dt,$$

which is strictly increasing on $\mathcal{D}_{\mathcal{J}_A}$ and satisfies $\mathcal{J}_A(0) = 0$. In particular, for any continuity point $a \geq 0$ of \mathcal{J}_A , we have

$$(2.10) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left(A_n \geq an^2 \right) = -\mathcal{J}_A(a).$$

The set of minimizers in (2.9), denoted by $H_A(a)$, is such that for any $\varepsilon > 0$,

$$(2.11) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{0 \leq k \leq n} \left| \frac{S_k}{n} - h(k/n) \right| \leq \varepsilon \text{ for some } h \in H_A(a) \mid A_n \geq an^2 \right) = 1, \\ a \in \text{int}(\mathcal{D}_{\mathcal{J}_A}).$$

(2) Suppose that the distribution of X_1 is rotationally invariant. Then $\mathcal{J}_A(a) = \underline{I}(\sqrt{2\pi a})$; equality (2.10) is valid for every $a \geq 0$; and for any $a \in [0, r_{max}^2/(2\pi))$,

$$(2.12) \quad H_A(a) = \left\{ \sqrt{\frac{2a}{\pi}} (\cos(\pm\pi t + \alpha) - \cos \alpha, \sin(\pm\pi t + \alpha) - \sin \alpha) \right\}_{\alpha \in \mathbb{R}}.$$

Thus, for rotationally invariant distributions, large deviations of the area are attained on the trajectories that asymptotically align into half-circles and move with constant speed. Note that for such distributions, \underline{I} is convex.

We will refer to the elements of the sets $H_A(a)$ as the *optimal trajectories* (for the area).

Remark 2.12. — Assume that the covariance matrix Σ of X_1 is non-degenerate, satisfies $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$, and $\Sigma^{-1/2}X_1$ has a rotationally invariant distribution, whose rate function we denote by I_1 . Then $(\Sigma^{-1/2}S_k)_{k \geq 1}$ is a random walk with a rotationally invariant distribution of increments, and the area of its convex hull satisfies

$$\text{area} \left(\text{conv} \left(\Sigma^{-1/2}S_1, \dots, \Sigma^{-1/2}S_n \right) \right) = (\det \Sigma)^{-1/2} A_n.$$

Hence by Theorem 2.11 we have $\mathcal{J}_A(a) = \underline{I}_1(\sqrt{2\pi a/(\det \Sigma)^{1/2}})$. To rewrite \mathcal{J}_A in terms of I , note that $\underline{I}_1(|u|) = I_1(u) = I(\Sigma^{1/2}u)$ for $u \in \mathbb{R}^2$, where the last equality follows

by changing variables in the definition of I . This gives $I_1(r) = \underline{I}(r\sqrt{\lambda_1})$ for $r \geq 0$, where $0 < \lambda_2 \leq \lambda_1$ are the eigenvalues of Σ , hence $\mathcal{J}_A(a) = \underline{I}(\sqrt{2\pi a}(\lambda_1/\lambda_2)^{1/4})$.

Moreover, equality (2.12) remains valid for $a \in [0, \text{area}(\text{supp}(X_1))/(2\pi^2))$ if we multiply the factor $\sqrt{2a/\pi}$ in (2.12) by $(\det \Sigma)^{-1/2}\Sigma^{1/2}$. The optimal trajectories are halves of the ellipse $\text{supp}(X_1)$ divided by the lines passing through its centre.

For random walks with a shifted (i.e. $\mu \neq 0$) rotationally invariant distribution of increments, the optimal trajectories for the area are not universal, unlike those for the perimeter. We were able to solve only the Gaussian case. We apply the same approach as in the proof of Theorem 2.11. In fact, since Gaussian rate functions are quadratic, computation of the rate function for the area reduces to finding a planar curve of fixed length and fixed endpoints that maximizes the area of its convex hull. Pach [Pac78] proved that such a curve is a circular arc, as in the Dido problem with fixed endpoints. The corresponding isoperimetric inequality yields the following LDP.

Let us denote by u^\perp a vector $u \in \mathbb{R}^2$ rotated $\pi/2$ counterclockwise about the origin.

THEOREM 2.13. — *Suppose that X_1 has a shifted standard Gaussian(μ, Id) distribution on \mathbb{R}^2 with a non-zero mean μ . Then*

$$(2.13) \quad \mathcal{J}_A(a) = 4a\varphi - \frac{1}{2}|\mu|^2 \tan^2 \varphi, \quad a \geq 0,$$

where $\varphi \in [0, \pi/2)$ is the unique solution to

$$\frac{2\varphi - \sin 2\varphi}{8\varphi^2 \cos^2 \varphi} = \frac{a}{|\mu|^2};$$

equality (2.10) is valid for every $a \geq 0$; and in the basis μ, μ^\perp , the set of optimal trajectories is

$$(2.14) \quad H_A(a) = \left\{ \frac{1}{2\varphi \cos \varphi} (\sin(2\varphi t - \varphi) + \sin \varphi, \pm \cos(2\varphi t - \varphi) \mp \cos \varphi) \right\}, \quad a > 0.$$

Thus, large deviations of the area of the convex hull for random walks with shifted standard Gaussian increments are attained on the trajectories that asymptotically align into either of the two μ -axially symmetric circular arcs of radius $\frac{|\mu|}{2\varphi \cos \varphi}$ and angle 2φ starting at the origin and ending on the μ -axis. The radius is defined so that for either of the two limit curves, the orthogonal projections of their velocities to the direction of μ at times 0 and 1 both equal μ . Note that the shifted standard Gaussian rate function $I(v) = \frac{1}{2}|v - \mu|^2$ is not constant on the velocity of the optimal trajectories (which move with constant speed), contrasting the results of Theorems 2.3 and 2.11.

It is easy to show that the asymptotics in Theorem 2.11 for the Gaussian case appears as the limit case of Theorem 2.13 as $|\mu| \rightarrow 0$ with a fixed a : since $\varphi \rightarrow \pi/2$, the radius tends to $\sqrt{2a/\pi}$ and the right-hand side of (2.13) tends to πa , which is $\underline{I}(\sqrt{2\pi a})$ for the standard Gaussian distribution.

Remark 2.14. — Using the same argument as in Remark 2.12 above, we can easily check that for X_1 following any Gaussian distribution with a non-degenerate covariance matrix Σ and non-zero drift μ , it holds that

$$\mathcal{J}_A(a) = \frac{4a\varphi}{\sqrt{\det \Sigma}} - \frac{1}{2} (\mu^\top \Sigma \mu) \tan^2 \varphi, \text{ where } \frac{2\varphi - \sin 2\varphi}{8\varphi^2 \cos^2 \varphi} = \frac{a}{\sqrt{\det \Sigma} \cdot \mu^\top \Sigma \mu}, \quad a \geq 0.$$

With this uniquely defined φ , the set of optimal trajectories is given by (2.14) taken in the basis $\mu, (\Sigma^{-1/2}\mu)^\perp$. In this general case the optimal limit shapes are elliptic arcs starting at the origin and ending on the μ -axis.

For completeness of exposition, we consider shifted degenerate Gaussian distributions. Since these arise as the limit case of non-degenerate Gaussian distributions, we can use Remark 2.14 to get the following result, which we present here without a proof.

PROPOSITION 2.15. — *Suppose that $X_1 \stackrel{d}{=} (\mu_1, \mu_2 + \sigma Y_1)$, where $\mu = (\mu_1, \mu_2)$, μ_1 and σ are non-zero, and Y_1 is a standard Gaussian random variable. Then $\mathcal{J}_A(a) = 6a^2\mu_1^{-2}\sigma^{-2}$ for $a \geq 0$; (2.10) is valid for every $a \geq 0$; and the optimal trajectories are the parabolas*

$$H_A(a) = \left\{ \mu t \pm 6a\mu_1^{-1} (0, t - t^2) \right\}, \quad a > 0.$$

In the case $\mu_1 = 1$, the proposition describes large deviations for the area of the convex hull of the graph of one-dimensional random walk with $\text{Gaussian}(\mu_2, \sigma^2)$ increments. The assumptions of Proposition 2.15 ensure that the distribution of X_1 is not supported on the line $\mu\mathbb{R}$ passing through the origin.

2.4. Convex hulls of Lévy processes

The above results on convex hulls of random walks, which have increments in discrete time, have the following counterparts in continuous time.

Assume now that $(S_t)_{t \geq 0}$ is a *Lévy process* on the plane, that is a stochastic process with stationary independent increments and càdlàg trajectories (i.e. right-continuous and having left limits) taking values in \mathbb{R}^2 . Then $(S_t)_{t \in \mathbb{N}}$ is a random walk. Conversely, every random walk with an *infinitely divisible* distribution of increments (that for every $n \in \mathbb{N}$ is the n -fold convolution of some distribution) can be regarded as such time-discretization of a Lévy process $(S_t)_{t \geq 0}$; see Bertoin [Ber96, Theorem I.1].

Consider the convex hull $C_T := \text{conv}(\{S_t\}_{0 \leq t \leq T})$ for $T > 0$. Its perimeter $P_T := P(C_T)$ and area $A_T := A(C_T)$ are random variables (see the Appendix).

The following general result extends Theorems 2.3(1) and 2.11(1).

THEOREM 2.16. — *Assume that $(S_t)_{t \geq 0}$ is a Lévy process on the plane such that $S_1 = X_1$ and $\mathcal{D}_C = \mathbb{R}^2$. Then the variables $(P_T/(2T))_{T > 0}$ and $(A_T/T^2)_{T > 0}$ satisfy the LDP's in \mathbb{R} (as $T \rightarrow \infty$) with speed T and the respective rate functions \mathcal{J}_P and \mathcal{J}_A (given in (2.4) and (2.9)).*

Remark 2.17. — Assume that $(S_t)_{t \geq 0}$ is a planar Brownian motion starting at zero (which is the only Lévy process with continuous trajectories). By Corollary 2.10

we have $\mathcal{J}_P = \underline{I}$, and \underline{I} can be found using the method of Lagrange multipliers as in the proof of Proposition 2.9(a) (where we essentially found a similar quantity \bar{K}). \mathcal{J}_A is given explicitly in Theorem 2.13, Proposition 2.15, and Remarks 2.12, 2.14, which together cover all possible values of the drift μ and the covariance Σ of S_1 .

Moreover, the limit shape results (2.7), (2.8), (2.11) remain valid with k and n considered as positive reals (instead of integers, as before) and P_n, A_n replaced respectively by $\mathbf{P}_n, \mathbf{A}_n$; this describes the optimal trajectories of the Brownian motion that result in the large deviations of \mathbf{P}_n and \mathbf{A}_n . Indeed, the proofs of Theorems 2.3 and 2.11 can be carried over without any changes if instead of Mogulskii's LDP (see Section 4.1) for trajectories of random walks we apply Schilder's LDP ([DZ10, Theorem 5.2.3]) for trajectories of a Brownian motion in \mathbb{R}^d (this result is stated in [DZ10] for a standard Brownian motion but we can convert it into the LDP for S applying the contraction principle to the mapping $f \mapsto \Sigma^{1/2}f(t) + \mu t$ for $f \in AC_0[0, 1]$).

We will prove Theorem 2.16 by reduction to the random walks case, showing that the trajectory of $(S_t)_{0 \leq t \leq T}$ stays close to that of $(S_t)_{t \in \{0, 1, \dots, [T]\}}$. This also allows one to extend our simplest limit shape result (2.7) to general Lévy processes. It is also tempting to provide counterparts to (2.8) and (2.11). However, it appears that arguing by reduction to random walks would require additional technical assumptions. Therefore, we do not give any results in this direction. Note that we cannot directly prove counterparts to (2.7), (2.8), (2.11), as we did for Brownian motions in Remark 2.17 above, since we are not aware of any LDP for general Lévy processes appropriate for the purpose.

2.5. Further extensions

2.5.1. Higher dimensions

One can further consider large deviations of surface area, volume, etc. for convex hulls of random walks in higher dimensions. The expected values of these quantities are available through the explicit formulas of [VZ18, Section 4] which generalize (1.1). However, currently we cannot obtain any progress even for rotationally invariant distributions of increments. In fact, according to Tilli [Til10], the problem of finding the shape of a curve in \mathbb{R}^d , where $d \geq 3$, of unit length that maximizes the volume of its convex hull is yet solved only in the class of curves convex in the sense of Schoenberg (i.e., those that intersect no hyperplane at more than d points) and there is no complete solution. Croft et al. [CFG91, Problem A28] mention that there are no results on the similar problem of maximizing the surface area, and we are unaware of any progress in this direction.

Remark 2.18. — On the other hand, our results for the perimeter of planar random walks can be easily extended for *mean width* of convex hulls in higher dimensions, defined in (A.1). A closely related quantity is the *first intrinsic volume* of the convex hull, which equals (see [SW08, Eq. (14.7)]) the mean width divided by $\frac{2v_{d-1}}{dv_d}$, which is mean width of a unit segment, where v_d denotes volume of a unit ball in \mathbb{R}^d .

Assume now that S_n is a random walk in \mathbb{R}^d satisfying $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^d$. Denote by W_n and V_n the mean width and first intrinsic volume, respectively, of the convex hull C_n . The Spitzer–Widom formula (1.1) remains valid (see [VZ18, Corollary 3]) in any dimension if we replace the perimeter P_n of the convex hull C_n by its doubled first intrinsic volume $2V_n$. Accordingly, our Theorem 2.3 remains valid if we replace P_n by $\frac{dv_d}{v_{d-1}}W_n = 2V_n$ so the probabilities change to $\mathbb{P}(W_n \geq \frac{2v_{d-1}}{dv_d}xn)$ and $\mathbb{P}(W_n \leq \frac{2v_{d-1}}{dv_d}xn)$ or, equivalently, more elegant expressions $\mathbb{P}(V_n \geq xn)$ and $\mathbb{P}(V_n \leq xn)$. The only difference in the proof is that Remark A.3 in the Appendix should be used instead of Corollary A.2.

2.5.2. Weaker exponential moments assumptions

The Cramér moment assumption $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$ for the increments is a standard minimal requirement to work with large deviations of random walks. However, in the case $\mathcal{D}_{\mathcal{L}} \neq \mathbb{R}^2$ we have to regard trajectories of the walk as random elements of the space of functions of bounded variation (equipped with a certain Skorokhod-type topology). Essentially, this is due to the fact that the rate function I is not super-linear at infinity, and in particular, the infima in (2.4) and (2.9) may not be attained on absolutely continuous functions.

The only available large deviations result for such trajectories is the non-standard LDP by Borovkov and Mogulskii [BM13]. It can be applied to our problems using the contraction principle by Vysotsky [Vys21a], which yields LDP’s for the perimeter and the area in the case $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$; see Proposition 4.1 in Section 4.4. The rate functions there are rather complicated but remarkably, they are exactly the same as in the main case $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$ when \underline{I} is convex (for the perimeter) or the distribution of X_1 is rotationally invariant (for the area). We do not identify the optimal trajectories in Proposition 4.1 since there are too many cases to analyse.

3. Properties of the radial minimum rate function \underline{I}

3.1. Basic facts from convex analysis

Suppose that $F : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ is any function with a non-empty effective domain \mathcal{D}_F .

- By [Roc70, Theorem 12.2 and Corollary 12.1.1]) it holds that

$$(3.1) \quad F^{**} = \text{cl}(\text{conv } F),$$

where $\text{cl}(\cdot)$ denotes the *closure* of a function, that is the function with the epigraph $\text{cl}(\text{epi}(\cdot))$. Recall that F is lower semi-continuous iff $F = \text{cl } F$, i.e. its epigraph $\text{epi } F$ is closed in $(\mathbb{R} \cup \{+\infty\}) \times \mathbb{R}^d$ ([Roc70, Theorem 7.1]). Thus the Legendre–Fenchel transform is an involution on the set of lower semi-continuous convex functions.

- The *Fenchel inequality* $F(u) + F^*(v) \geq u \cdot v$, which holds for any $u, v \in \mathbb{R}^d$, immediately follows from the definition of convex conjugation.

- If the function F is convex, then it is continuous on $\text{rint}(\mathcal{D}_F)$ ([Roc70, Theorem 10.1]) and if, in addition, F is lower semi-continuous, then it is continuous on

every closed interval contained in \mathcal{D}_F ([Roc70, Corollary 7.5.1]). If F is convex, finite, and differentiable on an open convex set $C \subset \mathbb{R}^d$, then F continuously differentiable on C ([Roc70, Corollary 25.5.1]). If $C = \mathbb{R}^d$, then F^* is *strictly convex* on $\text{rint}(\mathcal{D}_{F^*})$, which means that F^* is linear on no line segment with the endpoints in $\text{rint}(\mathcal{D}_{F^*})$ ([Roc70, Theorem 26.3]).

• Suppose that $d = 1$ and F is convex. Then

$$(3.2) \quad \{v \in \text{int}(\mathcal{D}_{F^*}) : F^* \text{ is affine on } [v - \varepsilon, v + \varepsilon] \text{ for some } \varepsilon > 0\} \\ = \text{int}(\text{conv}(\text{im}(F'))) \setminus \text{cl}(\text{im}(F')),$$

where, recall, $\text{im}(\cdot)$ denotes the image of a function. Thus, kinks of convex functions correspond to affine segments of their convex conjugates, and vice versa.

In order to prove this, let v belong to the set in the r.h.s. of (3.2). Since the function F' is non-decreasing on its domain and $v \notin \text{cl}(\text{im}(F'))$, there is a unique real u such that

$$(3.3) \quad \inf(\text{im}(F')) < F'_-(u) < v < F'_+(u) < \sup(\text{im}(F')).$$

Then it is easy to see that F^* is affine on $[F'_-(u), F'_+(u)]$ with slope u , which in particular implies that v belongs to the set in the l.h.s. of (3.2). For the reverse inclusion, if v belongs to the set in the l.h.s. of (3.2), which is open, then $v \in \text{int}(\text{conv}(\text{im}(F')))$ by $\mathcal{D}_{F^*} = \text{cl}(\text{conv}(\text{im}(F')))$. By taking the Legendre–Fenchel transform of F^* and using (3.1), which gives $F^{**} = F$ on $\text{int}(\mathcal{D}_F)$, we see that (3.3) holds true with $u = (F^*)'(v)$. Hence $v \notin \text{cl}(\text{im}(F'))$, and thus v belongs to the set in the r.h.s. of (3.2).

3.2. Basic properties of the radial minimum function \underline{I}

Example 3.1 (Discontinuous \underline{I}). — The function \underline{I} is not necessarily continuous on $[|\mu|, r_{max}]$: it is easy to check that if $\mathbb{P}(X_1 = (1, 0)) = 3/4$ and $\mathbb{P}(X_1 = (-2, 0)) = 1/4$, then \underline{I} has a jump at $r = 1$. It is also possible to show that \underline{I} is discontinuous for the “truly” two-dimensional distribution that is a mixture of the above two-atomic distribution and the uniform distribution on the disk $\{u : |u| \leq 1\}$.

It is not clear if \underline{I} can be discontinuous for zero mean distributions.

Proof of Lemma 2.1.

(a) This follows from (2.1) and the fact that $\mathcal{D}_{\underline{I}}$ is convex.

(b) Clearly, $\underline{I}(|\mu|) = 0$ by $I(\mu) = 0$ and $I \geq 0$. We claim that for any direction $\ell \in \mathbb{S}^{d-1}$, the function $I_\ell(t) := I(\mu + t\ell)$ is strictly increasing for $t \geq 0$ while it stays finite. This implies that \underline{I} strictly decreases on $[r_{min}, |\mu|]$ and strictly increases on $[|\mu|, r_{max}]$, since the line segment that joins μ with a point of minimum of I over the sphere $r\mathbb{S}^{d-1}$ always intersects the sphere $r'\mathbb{S}^{d-1}$ if $0 \leq r < r' < |\mu|$ or $|\mu| < r' < r$.

To prove the claim, we can assume without loss of generality that $\mu = 0$ since the rate function of $X_1 - \mu$ is $I(v - \mu)$ and the Laplace transforms of X_1 and $X_1 - \mu$ have the same effective domains. Since I_ℓ is a convex function with minimum at $t = 0$, it can cease to be strictly increasing only if it stays zero in a neighbourhood of 0. If $\ell \cdot X_1 = 0$ a.s., then there is nothing to prove since $I_\ell(t) = +\infty$ for $t > 0$

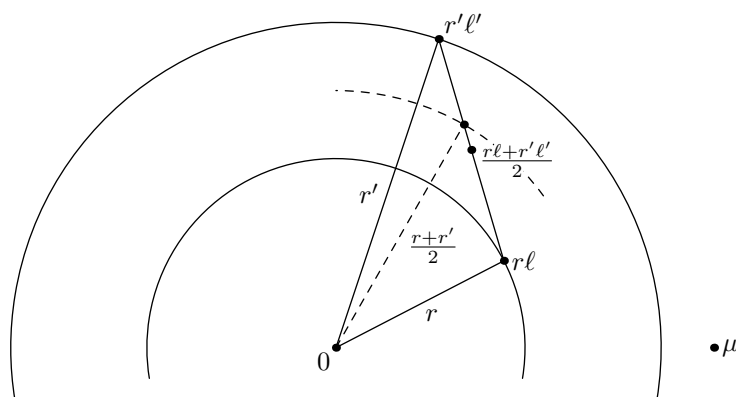


Figure 3.1.

by (2.1), otherwise by the criterion of equality in Hölder’s inequality, the function $a \mapsto \log \mathbb{E}e^{al \cdot X_1}$, where $a \in \mathbb{R}$, is strictly convex on its effective domain. Since the interior of this domain contains 0 by the assumption $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$, it is easy to see that $I(t\ell) \geq \sup_{a \in \mathbb{R}} (al \cdot t\ell - \log \mathbb{E}e^{al \cdot X_1}) > 0$ for $t \neq 0$. This proves the claim.

Furthermore, the function \underline{I} is convex on $[r_{\min}, |\mu|]$ since for any $r_{\min} \leq r < r' \leq |\mu|$ and $\ell \in \underline{\Delta}_r, \ell' \in \underline{\Delta}_{r'}$, one has

$$\underline{I}(r) + \underline{I}(r') = I(r\ell) + I(r'\ell') \geq 2I\left(\frac{r\ell + r'\ell'}{2}\right) \geq 2\underline{I}\left(\left|\frac{r\ell + r'\ell'}{2}\right|\right) \geq 2\underline{I}\left(\frac{r + r'}{2}\right).$$

Here we used the triangle inequality and the fact that \underline{I} decreases on $[0, |\mu|]$, see Figure 3.2 for a geometric explanation in the planar case. If $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^d$, then I is strictly convex on its effective domain, thus the first inequality is strict, hence \underline{I} is strictly convex on $[r_{\min}, |\mu|]$.

The lower semi-continuity of \underline{I} easily follows from that of I using a simple compactness argument.

(c) By Part (b), \underline{I} is lower semi-continuous and increasing, and hence left-continuous, on $[|\mu|, r_{\max}]$. If \underline{I} is discontinuous at an $x \in [|\mu|, r_{\max}]$, then it must be $\underline{I}(x) < \infty$, otherwise there is a contradiction with the left-continuity of \underline{I} . For any $\ell \in \underline{\Delta}_x$, consider the hyperplane L passing through $x\ell$ and orthogonal to ℓ . Assume that there is an $v \in L \cap \mathcal{D}_I$ that is distinct from $x\ell$. Then $|(1 - \varepsilon)x\ell + \varepsilon v| > x$ for every $\varepsilon > 0$, hence

$$\underline{I}(x+) \leq \lim_{\varepsilon \rightarrow 0+} I((1 - \varepsilon)x\ell + \varepsilon v) = I(x\ell) = \underline{I}(x),$$

where the first equality holds since the convex lower semi-continuous function I is continuous on the line segment $[x\ell, v] \subset \mathcal{D}_I$ ([Roc70, Corollary 7.5.1]). Thus, \underline{I} is continuous at x , which is a contradiction. Therefore, $L \cap \mathcal{D}_I = x\ell$, meaning that $x\ell$ is an exposed point of \mathcal{D}_I . It remains to check that $I(x\ell) = -\log \mathbb{P}(X_1 = x\ell)$.

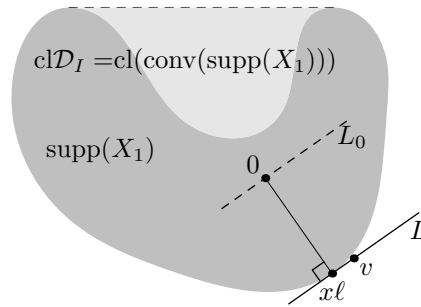


Figure 3.2.

We have $\ell \cdot X_1 \leq x$ a.s., where the inequality is strict unless $X_1 = x\ell$. Denote by L_0 the hyperplane passing through 0 and parallel to L ; see Figure 3.2. Let us identify \mathbb{R}^d with $L_0 \oplus \mathbb{R}\ell$. For any $u_1 \in L_0$ such that $\mathbb{E}e^{(u_1+u_2\ell)\cdot X_1} < \infty$ for some real u_2 , we have

$$\begin{aligned}
 (3.4) \quad \sup_{u_2 \in \mathbb{R}} (u_2x - \log \mathbb{E}e^{(u_1+u_2\ell)\cdot X_1}) &= -\log \left(\inf_{u_2 \in \mathbb{R}} \mathbb{E}e^{u_1 \cdot X_1 + u_2(\ell \cdot X_1 - x)} \right) \\
 &= -\log \mathbb{E} \left[e^{u_1 \cdot X_1} \mathbb{1}_{\{X_1 = x\ell\}} \right] = -\log \mathbb{P}(X_1 = x\ell),
 \end{aligned}$$

with the second equality following from the dominated convergence theorem using that the random variables in the $\inf \mathbb{E}$ term decrease point-wisely in u_2 since $\ell \cdot X_1 \leq x$ a.s. These equalities hold true e.g. for $u_1 = 0$. This yields the required equality

$$I(x\ell) = \sup_{u_1 \in L_0} \sup_{u_2 \in \mathbb{R}} \left((u_1 + u_2\ell) \cdot x\ell - \log \mathbb{E}e^{(u_1+u_2\ell)\cdot X_1} \right) = -\log \mathbb{P}(X_1 = x\ell),$$

where the last equality holds because no $u_1 \in L_0$ such that $\mathbb{E}e^{(u_1+u_2\ell)\cdot X_1} = \infty$ for every $u_2 \in \mathbb{R}$ contributes to the first supremum since for such u_1 the l.h.s. of the first line in (3.4) is $-\infty$.

Finally, we have $\mathbb{P}(X_1 = x\ell) > 0$ by $x\ell \in \mathcal{D}_I$.

(d) Suppose that for an $r \in (r_{min}, |\mu|]$, there are two distinct elements ℓ, ℓ' in $\underline{\Delta}_r$. By convexity of I , it holds that $\underline{I}(r|\ell + \ell'|/2) \leq I(r(\ell + \ell')/2) \leq \underline{I}(r)$, which is a contradiction by Part (b) since \underline{I} is strictly decreasing on $(r_{min}, |\mu|]$ and $|\ell + \ell'| < 2$. \square

3.3. Radial maxima and minima of conjugate convex functions

Let us prove the following statement, which may be known in convex analysis but we found no references. It is stronger than Proposition 2.6 since the Laplace transform of a distribution is lower semi-continuous by Fatou’s lemma. In particular, it applies to distributions with Laplace transform finite only in a neighbourhood of zero.

PROPOSITION 3.2. — *Let $F : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$, where $d \geq 1$, be any lower semi-continuous convex function differentiable at 0 and such that $p_{min} := \inf\{|u| : F(u) = \infty\} > 0$, where $\inf_\emptyset = \infty$ by convention. Put $m := \nabla F(0)$ and define \bar{F} and \underline{F}^* as in (2.2). Then*

(a) \bar{F} is an increasing convex function on $[0, \infty)$ satisfying $\bar{F}'_+(0) = |m|$ and

$$\text{conv } \underline{F}^* = (\bar{F})^* \text{ on } [|m|, \infty);$$

(b) If $r \in \text{cl}(\text{im}(\bar{F}'))$ (and $r \geq |m|$), then $\underline{F}^*(r) = \text{conv}(\underline{F}^*)(r) < \infty$.

If additionally F is differentiable on $\{u : |u| < p_{\min}\}$, and $\bar{\Lambda}_p$ and $\underline{\Lambda}_r$ are defined for F and F^* as in (2.3), then

(c) For any $p \in (0, p_{\min})$, the one-sided derivatives satisfy

$$\bar{F}'_+(p) = \max_{\ell \in \bar{\Lambda}_p} |\nabla F(p\ell)| \quad \text{and} \quad \bar{F}'_-(p) = \min_{\ell \in \bar{\Lambda}_p} |\nabla F(p\ell)|.$$

(d) If $p \in (0, p_{\min})$ and $r \geq |m|$ are such that $\bar{F}'(p) = r$, then $\underline{\Lambda}_r = \bar{\Lambda}_p$.

The following corollaries to Proposition 3.2 will easily imply those to Proposition 2.6.

COROLLARY 3.3. — *If \bar{F} is differentiable on $(0, p_{\min})$, then \underline{F}^* is strictly convex on $[|m|, (\bar{F})'_-(p_{\min})]$.*

COROLLARY 3.4. — *\bar{F} is differentiable on $(0, p_{\min})$ if F is differentiable on $\{u : |u| < p_{\min}\}$ and there exists a continuous mapping $\ell : (0, p_{\min}) \rightarrow \mathbb{S}^{d-1}$ such that $\ell(p) \in \bar{\Lambda}_p$ for any $p \in (0, p_{\min})$.*

There are few ways to prove Part (a) of Proposition 3.2, using geometric or analytic approaches. The current simple proof is due to Fedor Petrov.

Proof of Proposition 3.2.

(a) We have $\bar{F}(p) = \sup_{\ell \in \mathbb{S}^{d-1}} F(p\ell)$, where $p \geq 0$, hence \bar{F} is convex as a maximum of convex functions $F_\ell(\cdot) := F(\cdot\ell)$. Furthermore, the convex function F attains its maximum over any closed compact convex set on the boundary of the set. Therefore for any $0 \leq p < p' \leq p_{\min}$, we have

$$\bar{F}(p) = \max_{u: |u|=p} F(u) = \max_{u: |u|\leq p} F(u) \leq \sup_{u: |u|\leq p'} F(u) = \bar{F}(p')$$

(where the supremum may not be attained if $p' = p_{\min}$). Hence \bar{F} is increasing on $[0, \infty)$ since $\bar{F}(p) = \infty$ for $p > p_{\min}$. The right derivative of \bar{F} at 0 clearly satisfies $\bar{F}'_+(0) = |m|$.

It remains to prove that $\text{conv}(\underline{F}^*) = (\bar{F})^*$ on $[|m|, \infty)$. We first claim that

$$(3.5) \quad (\underline{F}^*)^*(p) = \bar{F}(p), \quad p \geq 0.$$

In fact, by the definition, we have $\underline{F}^*(r) = \infty$ for $r < 0$, hence for every real p ,

$$\begin{aligned} (\underline{F}^*)^*(p) &= \sup_{r \geq 0} (pr - \underline{F}^*(r)) = \sup_{r \geq 0} \left(pr - \inf_{\ell \in \mathbb{S}^{d-1}} F^*(r\ell) \right) \\ &= \sup_{r \geq 0, \ell \in \mathbb{S}^{d-1}} (pr - F^*(r\ell)) = \sup_{v \in \mathbb{R}^d} (p|v| - F^*(v)). \end{aligned}$$

On the other hand, by the assumptions, F is convex, lower semi-continuous, and has non-empty effective domain, hence $F = F^{**}$ holds by (3.1). Then (3.5) follows since for $p \geq 0$,

$$\bar{F}(p) = \sup_{\ell \in \mathbb{S}^{d-1}} F(p\ell) = \sup_{\ell \in \mathbb{S}^{d-1}} \sup_{v \in \mathbb{R}^d} (p\ell \cdot v - F^*(v)) = \sup_{v \in \mathbb{R}^d} (p|v| - F^*(v)).$$

By the definition, we have $\bar{F}(p) = \infty$ for $p < 0$, hence the Legendre–Fenchel transform of \bar{F} is fully defined by the values of \bar{F} on $[0, \infty)$. Likewise, the Legendre–Fenchel transform of $(\underline{F}^*)^*$ restricted to $[|m|, \infty)$ is defined by the values of $(\underline{F}^*)^*$ on $[0, \infty)$. In fact, the function $p \mapsto pr - (\underline{F}^*)^*(p)$ is increasing on $(-\infty, 0]$ for any $r \geq |m|$ since $(\underline{F}^*)^*$ is a convex function, whose right derivative increases on its domain and its value at 0 equals that of \bar{F} by (3.5), while we already proved that $\bar{F}'_+(0) = |m|$. Therefore, (3.5) implies $(\underline{F}^*)^{**} = (\bar{F})^*$ on $[|m|, \infty)$, hence $\text{cl}(\text{conv}(\underline{F}^*)) = (\bar{F})^*$ on $[|m|, \infty)$ by (3.1).

It remains to remove the closure operation cl from the last equality. It suffices to show that $\text{conv}(\underline{F}^*)$ is lower semi-continuous at the boundary points of its effective domain, which coincides with that of \underline{F}^* . By a simple compactness argument it follows from the lower semi-continuity of \underline{F}^* that \underline{F}^* is also lower semi-continuous. Every point $x \in \partial(\mathcal{D}_{\underline{F}^*})$ has a neighbourhood U such that the convex function $\text{conv}(\underline{F}^*)$ is either strictly increasing, strictly decreasing, or constant on $U \cap \mathcal{D}_{\underline{F}^*}$. From the definition of the largest convex minorant, it follows that $\text{conv}(\underline{F}^*)(x) = \underline{F}^*(x)$. This equality, combined with the property of lower semi-continuity of \underline{F}^* at x , implies the same property for $\text{conv}(\underline{F}^*)$ by a simple consideration of the three cases mentioned above.

(b) It is easy to see that the non-negative function $\underline{F}^* - \text{conv}(\underline{F}^*)$, which we define to be zero outside $\mathcal{D}_{\underline{F}^*}$, is lower semi-continuous. In fact, this property holds at the points of $\text{int}(\mathcal{D}_{\underline{F}^*})$ by continuity of $\text{conv}(\underline{F}^*)$ and lower semi-continuity of \underline{F}^* , which we showed above in the proof of Part (a). At the points of $\partial(\mathcal{D}_{\underline{F}^*})$, this is true by non-negativity and the fact that $\underline{F}^* = \text{conv}(\underline{F}^*)$ on $\partial(\mathcal{D}_{\underline{F}^*})$. Hence the set

$$\{r \geq |m| : \underline{F}^*(r) = \text{conv}(\underline{F}^*)(r)\} = \{r \geq |m| : \underline{F}^*(r) - \text{conv}(\underline{F}^*)(r) \leq 0\}$$

is closed as a sub-level set of a lower semi-continuous function. Therefore, if $\underline{F}^*(r) > \text{conv}(\underline{F}^*)(r)$ for an $r \in (|m|, \sup \mathcal{D}_{\underline{F}^*})$, then this inequality also holds on an open interval (r_1, r_2) that contains r , on which $\text{conv}(\underline{F}^*)$ must be affine. Since $\text{conv}(\underline{F}^*) = (\bar{F})^*$ on $[|m|, \infty)$ by Part (a), we conclude that $(\bar{F})^*$ is affine on $[r_1, r_2]$. As we explained in Section 3.1, this yields that $r \notin \text{cl}(\text{im}(\bar{F}'))$, which is a contradiction.

(c) The one-sided derivatives of \bar{F} exist by convexity of this function proven in Part (a). The set $\bar{\Lambda}_p := \text{argmax}_{\ell \in \mathbb{S}^{d-1}} F(p\ell)$ is well-defined since the convex function F is continuous on $\text{rint}(\mathcal{D}_F)$ and $p\mathbb{S}^{d-1}$ is a compact subset of $\text{rint}(\mathcal{D}_F)$ by $p \in (0, p_{\min})$.

For any $\ell \in \bar{\Lambda}_p$, the gradient $\nabla F(p\ell)$ is directed along ℓ since $p\ell$ is an extremal point of the function F over the sphere $p\mathbb{S}^{d-1}$ and F is differentiable, and hence continuously differentiable, on $\{u : |u| < p_{\min}\}$; see Section 3.1. Hence $|\nabla F(p\ell)| = F'_\ell(p)$ and by

$$\bar{F}'_+(p) = \lim_{\varepsilon \rightarrow 0^+} \varepsilon^{-1}(\bar{F}(p + \varepsilon) - \bar{F}(p)) \geq \lim_{\varepsilon \rightarrow 0^+} \varepsilon^{-1}(F((p + \varepsilon)\ell) - F(p\ell)) = F'_\ell(p),$$

we arrive at $\bar{F}'_+(p) \geq \max_{\ell \in \bar{\Lambda}_p} |\nabla F(p\ell)|$, where the r.h.s. accounts the fact that the function ∇F , which is continuous on $p\mathbb{S}^{d-1}$, attains its maximum on the compact set $p\bar{\Lambda}_p$.

Furthermore, since \mathbb{S}^{d-1} is compact, there exist two sequences $p_k \rightarrow p+$ and $\ell(k) \in \bar{\Lambda}_{p_k}$ such that $\ell(k) \rightarrow \ell$ for some $\ell \in \mathbb{S}^{d-1}$ as $k \rightarrow \infty$. Then necessarily $\ell \in \bar{\Lambda}_p$ since F and \bar{F} are continuous on some neighbourhoods of $p\mathbb{S}^{d-1}$ and p , respectively.

Finally,

$$\begin{aligned} \bar{F}(p_k) - \bar{F}(p) &= F(p_k \ell(k)) - F(p\ell) \\ &= (p_k \ell(k) - p\ell) \cdot (\nabla F(p\ell) + o(1)) \leq (p_k - p) (|\nabla F(p\ell)| + o(1)) \end{aligned}$$

as $k \rightarrow \infty$, and thus $\bar{F}'_+(p) \leq \max_{\ell \in \bar{\Lambda}_p} |\nabla F(p\ell)|$. This inequality, combined with the opposite one proven above, yields the equality required.

The argument for $\bar{F}'_-(p)$ is analogous.

(d) The set $\underline{\Lambda}_r := \operatorname{argmin}_{\ell \in \mathbb{S}^{d-1}} F^*(r\ell)$ is well-defined since F^* is lower semi-continuous. First check that $\bar{\Lambda}_p \subset \underline{\Lambda}_r$. For any $\ell \in \bar{\Lambda}_p$, $\nabla F(p\ell)$ is directed along ℓ , hence by Part (c) it holds that $\nabla F(p\ell) = r\ell$. Note that $F(u) \geq F(p\ell) + r\ell \cdot (u - p\ell)$ for any $u \in \mathbb{R}^d$ since the right-hand side of this inequality defines the support hyperplane to graph of F at the point $(p\ell, F(p\ell))$. Then

$$\begin{aligned} (3.6) \quad \underline{F}^*(r) &\leq F^*(r\ell) = \sup_{u \in \mathbb{R}^d} (r\ell \cdot u - F(u)) \\ &\leq \sup_{u \in \mathbb{R}^d} (r\ell \cdot u - F(p\ell) - r\ell \cdot (u - p\ell)) = r\ell \cdot p\ell - F(p\ell) = rp - \bar{F}(p). \end{aligned}$$

The concave function $q \mapsto rq - \bar{F}(q)$ attains its maximum at $q = p$ since by the assumption, it holds that $\bar{F}'(p) = r$. Then by Parts (a) and (b),

$$(3.7) \quad rp - \bar{F}(p) = (\bar{F})^*(r) = \operatorname{conv}(\underline{F}^*)(r) = \underline{F}^*(r),$$

and since the latter expression equals the first term in (3.6), we get $\ell \in \underline{\Lambda}_r(F^*)$.

It remains to prove the reverse inclusion $\underline{\Lambda}_r \subset \bar{\Lambda}_p$. Suppose that $\ell \in \underline{\Lambda}_r$. Combining the Fenchel inequality with (3.7), we obtain

$$\bar{F}(p) \geq F(p\ell) \geq r\ell \cdot p\ell - F^*(r\ell) = rp - \underline{F}^*(r) = \bar{F}(p),$$

which implies that $\ell \in \bar{\Lambda}_p$. □

Proof of Corollary 3.3. Since the function \bar{F} , which is convex on $[0, \infty)$ by Proposition 3.2(a), is assumed to be differentiable on $(0, p_{min})$, it is continuously differentiable there; see Section 3.1. Then by (3.2), $(\bar{F})^*$ is affine on no subinterval of

$$\left[\inf(\operatorname{im}(\bar{F}')), \sup(\operatorname{im}(\bar{F}')) \right] = \left[(\bar{F}')_+(0), (\bar{F}')_-(p_{min}) \right] = \left[|m|, (\bar{F}')_-(p_{min}) \right],$$

and therefore strictly convex there. So is the function \underline{F}^* , which equals $(\bar{F})^*$ on $[|m|, \infty)$ by Proposition 3.2(a) and 3.2(b). □

Proof of Corollary 3.4. Since the function \bar{F} is convex on $[0, \infty)$, its left and right derivatives satisfy ([Roc70, Theorem 24.1])

$$\bar{F}'_+(p-) = \bar{F}'_-(p) \leq \bar{F}'_+(p) = \bar{F}'_-(p+), \quad p \in (0, p_{min}).$$

On the other hand, we have $\bar{F}'_-(p) \leq |\nabla F(p\ell(p))| \leq \bar{F}'_+(p)$ by Proposition 3.2(c). The claim follows by combining these relations and using that $|\nabla F(p\ell(p))|$ is continuous on $(0, p_{min})$, which is true since $p\ell(p)$ is continuous on $(0, p_{min})$ and ∇F is continuous on $\operatorname{int}(\mathcal{D}_F)$; see Section 3.1. □

Proofs of Corollaries 2.7 and 2.8. We apply Corollaries 3.3 and 3.4 with K substituted for F . Since $\mathcal{D}_\mathcal{L} = \mathbb{R}^d$ by the assumption, we have $p_{min} = \infty$ by $K = \log \mathcal{L}$.

Then Corollary 2.8 follows from Corollary 3.4. Furthermore, it follows from Proposition 2.6(c) that $\lim_{p \rightarrow \infty} (\bar{K})'_-(p) = r_{max}$. Then \underline{I} is strictly convex on $[|\mu|, r_{max}]$ by Corollary 3.3, while \underline{I} is strictly convex on $[r_{min}, |\mu|]$ by Lemma 2.1 (b). Since I attains its minimum at μ and is continuous at μ except for the trivial case $X_1 = \mu$ a.s., \underline{I} is strictly convex on the interval $[r_{min}, r_{max}]$, which contains $\mathcal{D}_{\underline{I}}$ by Lemma 2.1 (a). This proves Corollary 2.7, which is trivial in the remaining case $X_1 = \mu$ a.s. \square

3.4. Convexity of the radial minimum function \underline{I}

Here we prove that \underline{I} is convex for the distributions described in Proposition 2.9.

Proof of Proposition 2.9.

(a) Denote $r := \text{rank } A$. We assume that $r \geq 1$, otherwise the claim is trivial. The $d \times k$ matrix A admits a singular value decomposition $A = UDV^\top$, where D is an $r \times r$ diagonal matrix whose diagonal entries are non-zero singular values of A (i.e., the square roots of non-zero eigenvalues of AA^\top), and U is $d \times r$ matrix and V is a $k \times r$ matrix such that both $U^\top U$ and $V^\top V$ are the unit $r \times r$ matrices.

Put $L := U\mathbb{R}^r$. Then the assumption $AA^\top \mu = \|AA^\top\| \mu$ implies that $\mu \in L$. Furthermore, it is easy to check that $(U^\top u_1, U^\top u_2) = (u_1, u_2)$ for any $u_1, u_2 \in L$ and $|U^\top u| < |u|$ for $u \in \mathbb{R}^d \setminus L$. Therefore, from the equalities

$$\begin{aligned} \|AA^\top\| &= \max_{u \in \mathbb{S}^{d-1}} |AA^\top u| &&= \max_{u \in \mathbb{S}^{d-1}} |UD^2U^\top u| \\ &= \max_{u \in \mathbb{S}^{d-1} \cap L} |UD^2U^\top u| &&= \max_{u' \in \mathbb{S}^{r-1}} |D^2 u'| = \|D\|^2, \end{aligned}$$

and the assumption $AA^\top \mu = \|AA^\top\| \mu$, we see that the vector $\mu' := U^\top \mu$ in \mathbb{R}^r satisfies $D\mu' = \sigma_1 \mu'$, where $\sigma_1 := \|D\|$ is the largest singular value of A .

Then for any $v \in L$, by $X_1 \in L$ a.s. we have

$$\begin{aligned} I(v) &= \sup_{u \in \mathbb{R}^d} (u \cdot v - \log \mathbb{E} e^{u \cdot X_1}) = \sup_{u \in L} (u \cdot (v - \mu) - \log \mathbb{E} e^{u \cdot AY_1}) \\ &= \sup_{u \in L} (U^\top u \cdot U^\top (v - \mu) - \log \mathbb{E} e^{DU^\top u \cdot V^\top Y_1}). \end{aligned}$$

Denote by J the rate function of the random vector $V^\top Y_1$ in \mathbb{R}^r . Let us use that $I(v) = +\infty$ for $v \in \mathbb{R}^d \setminus L$ (by $X_1 \in L$ a.s.) and change variables $u' = U^\top u$, $v' = U^\top v$, $u'' = Du'$ to get

$$\begin{aligned} \underline{I}(r) &= \min_{v \in r\mathbb{S}^{d-1} \cap L} I(v) = \min_{v' \in r\mathbb{S}^{r-1}} \sup_{u' \in \mathbb{R}^r} (u' \cdot (v' - \mu') - \log \mathbb{E} e^{Du' \cdot V^\top Y_1}) \\ &= \min_{v' \in r\mathbb{S}^{r-1}} \sup_{u'' \in \mathbb{R}^r} (D^{-1}u'' \cdot (v' - \mu') - \log \mathbb{E} e^{u'' \cdot V^\top Y_1}) \\ &= \min_{v' \in r\mathbb{S}^{r-1}} J(D^{-1}(v' - \mu')), \end{aligned}$$

where in the last equality we also used that D^{-1} is symmetric.

The distribution of $V^\top Y_1$ on \mathbb{R}^r is rotationally invariant since so is that of Y_1 on \mathbb{R}^k . Therefore, J is a radial function, hence

$$\underline{I}(r) = \underline{J} \left(\min_{v' \in r\mathbb{S}^{r-1}} |D^{-1}(v' - \mu')| \right) = \underline{J}(\sigma_1^{-1} |r - |\mu'||), \quad r \geq 0,$$

where $|\mu'| = |\mu|$ and we used that $D^{-1}\mu' = \sigma_1^{-1}\mu'$ and σ_1^{-1} is the smallest eigenvalue of D^{-1} . Hence \underline{I} is convex since so is \underline{J} and \underline{J} is increasing on $[0, \infty)$.

(b) For $d = 1$, \underline{I} is convex by Part (a), so we assume that $d \geq 2$. We will give a detailed treatment for illustrative purposes in the planar case and then proceed to higher dimensions.

(1) The planar case $d = 2$ with non-degenerate covariance matrix Σ of X_1 .

The cumulant generating function K of a Gaussian(μ, Σ) distribution is $K(u) = u^\top \mu + \frac{1}{2}u^\top \Sigma u$. By Corollary 2.7, it suffices to show that the radial maximum function $\bar{K}(p)$ is differentiable on $(0, \infty)$. Since \bar{K} is invariant under orthogonal transformations of \mathbb{R}^2 , without loss of generality we can assume that

$$K(x, y) = \frac{1}{2}a(x - x_0)^2 + \frac{1}{2}b(y - y_0)^2 + c,$$

where a and b are the eigenvalues of Σ , $c = -\frac{1}{2}ax_0^2 - \frac{1}{2}by_0^2$, $\mu = (-ax_0, -by_0)$, and $x_0, y_0 \geq 0$ by changing directions of the axes, if necessary. We can further assume that $a > b > 0$ and $x_0 + y_0 > 0$, since the cases $a = b$ and $x_0 = y_0 = 0$ are already covered by Part (a).

To prove that \bar{K} is differentiable, by Corollary 2.8 it suffices to show that there is a continuous path $\ell(p)$ on the unit sphere that belongs to $\bar{\Lambda}_p$ for every $p > 0$. Suppose that $\ell \in \bar{\Lambda}_p$, i.e. K attains its maximum over $p\mathbb{S}^1$ at the point $p\ell$. Then $\nabla K(p\ell) = t p\ell$ for some non-zero real t , that is $(a(x - x_0), b(y - y_0)) = (tx, ty)$. Equivalently,

$$(3.8) \quad (a - t)x = ax_0, \quad (b - t)y = by_0.$$

(a) The case $x_0, y_0 > 0$. The set $\bar{\Lambda}_p$ lies in the quadrant $\{(x, y) : x \leq 0, y \leq 0\}$ since $K(-|x|, -|y|) < K(x, y)$ for any pair (x, y) in the complement of the quadrant. Hence, because the right-hand sides of the equalities in (3.8) are strictly positive, we have $t > a$. Therefore, equalities (3.8) define the curve

$$(3.9) \quad h(t) := \left(\frac{ax_0}{a - t}, \frac{by_0}{b - t} \right), \quad t > a,$$

marked in bold in Figure 3.3(a). Note in passing that $h(t)$ is an arc of the Apollonian hyperbola for the ellipses that are contour lines of K ; see Glaeser et al. [GSO16, Section 9.3] for details.

Both coordinates of $h(t)$ are strictly decreasing and continuous in t , hence the function $t \mapsto |h(t)|$ is a strictly decreasing continuous bijection from (a, ∞) to $(0, \infty)$. Therefore, the curve in (3.9), obtained from the necessary condition (3.8) for a maximum, has a unique point of intersection with the circle $p\mathbb{S}^1$. This point must be the unique element of the non-empty set $\bar{\Lambda}_p$. Thus, $\ell(p) := h(|h|^{-1}(p))$ is the curve required.

(b) The cases $x_0 > 0, y_0 = 0$ and $x_0 = 0, y_0 > 0$. We consider them solely for the purpose of illustration since they will be covered below in Part (3.4) using a different general argument. Meanwhile, note in passing that here equalities (3.8) define two lines $x = \frac{ax_0}{a-b}$ and $y = -\frac{by_0}{a-b}$, which can be regarded as the limit shapes for the hyperbolas in (3.9).

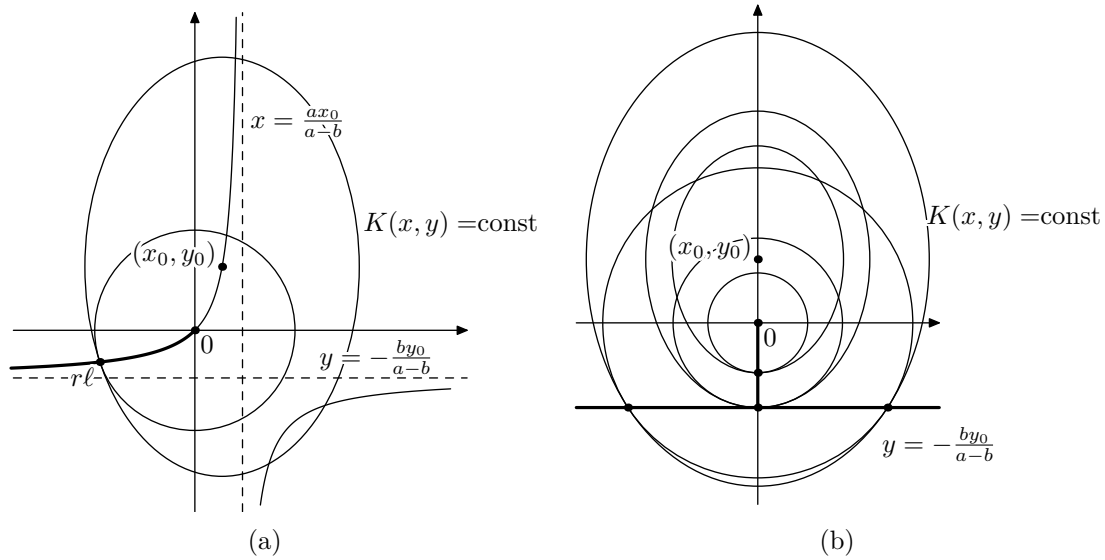


Figure 3.3.

It is easy to see that in the first case $y_0 = 0$, we have $\bar{\Lambda}_p = \{(-1, 0)\}$ for every $p > 0$, so $\ell(p) := (-1, 0)$; this situation is actually covered above in Part (a). In the second case $x_0 = 0$, from (3.8) we have $x = 0$ or $a = t$. Both solutions contribute to the answer – we have $\bar{\Lambda}_p = \{(0, -1)\}$ for $p \in (0, by_0/(a - b)]$, and $\bar{\Lambda}_p$ consists of two directions symmetric about the y -axis for $p > by_0/(a - b)$. The set $\cup_{p>0} p\bar{\Lambda}_p$ is marked in bold in Figure 3.3(b). Clearly, there is a continuous path $\ell(p)$ of directions such that $\ell(p) \in \bar{\Lambda}_p$ for every $p > 0$, as required.

(2) Arbitrary dimension $d \geq 2$ with non-degenerate Σ .

Take a basis of \mathbb{R}^d of eigenvectors of Σ , where the coordinates μ_i of μ are non-positive. If all μ_i 's are strictly negative, we argue exactly as above, putting $\ell(p) := h(|h|^{-1}(p))$, where

$$h(t) := - \left(\frac{\mu_1}{\sigma_1^2 - t}, \dots, \frac{\mu_d}{\sigma_d^2 - t} \right), \quad t > \sigma_1^2,$$

and $\sigma_1^2 \geq \dots \geq \sigma_d^2 > 0$ are the eigenvalues of Σ . Note in passing that if $\sigma_1^2 = \dots = \sigma_d^2$, then $h(t)$ parametrizes the half-line emanating from 0 in the direction of μ .

If some coordinates of μ are zero, we proceed differently from our consideration in the planar case and prove the convexity of \underline{I} directly. The rate function of the Gaussian(μ, Σ) distribution is given by $I(v) = \frac{1}{2}(v - \mu)^\top \Sigma^{-1}(v - \mu)$. For any $\varepsilon > 0$, the Gaussian($\mu - \varepsilon e_d, \Sigma$) distribution, where $e_d := (1, \dots, 1)$, has the rate function $I_\varepsilon(v) := I(v + \varepsilon e_d)$. All coordinates of $\mu - \varepsilon e_d$ are strictly negative, hence each function $\underline{I}_\varepsilon$ is convex on $[0, \infty)$ as shown above. On the other hand, $I_\varepsilon \rightarrow I$ as $\varepsilon \rightarrow 0+$ uniformly on every compact subset of \mathbb{R}^d since I is continuous on \mathbb{R}^d , and hence locally uniformly continuous. Then $\underline{I}_\varepsilon(r) \rightarrow \underline{I}(r)$ for every $r \geq 0$, which implies that \underline{I} is convex on $[0, \infty)$, as required.

(3) Arbitrary dimension $d \geq 2$ with degenerate Σ .

Put $L := \Sigma \mathbb{R}^d$ and note that $\Sigma_L := \Sigma|_L$ is a bijection from L to L . The rate function of the Gaussian(μ, Σ) distribution with degenerate Σ is given by $I(v) = \frac{1}{2}(v - \mu)^\top \Sigma_L^{-1}(v - \mu)$ for $v \in \mu + L$ and $I(v) = +\infty$ for $v \notin \mu + L$.

For any $\varepsilon > 0$, let Σ_ε be the positive definite $d \times d$ matrix defined by $\Sigma_\varepsilon u = \Sigma u$ for $u \in L$ and $\Sigma_\varepsilon u = \varepsilon u$ for $u \in \ker \Sigma$. Let I_ε be the rate function of the Gaussian(μ, Σ_ε) distribution. We have $I_\varepsilon(v) = I(v)$ for $v \in \mu + L$ and $I_\varepsilon(v) \nearrow I(v)$ as $\varepsilon \rightarrow 0+$ for $v \notin \mu + L$. Since the matrix Σ_ε is non-degenerate, each function $\underline{I}_\varepsilon$ is convex on $[0, \infty)$ as shown above. To conclude that \underline{I} is convex on $[0, \infty)$, it remains to prove that $\underline{I}_\varepsilon(r) \rightarrow \underline{I}(r)$ for every $r \geq 0$.

Denote by v' the orthogonal projection on L of a $v \in \mathbb{R}^d$ and put $v'' := v - v'$. Then

$$I_\varepsilon(v) = \frac{1}{2}(v' - \mu')^\top \Sigma_\varepsilon^{-1}(v' - \mu') + \frac{1}{2}\varepsilon^{-1}|v'' - \mu''|^2 = I(v' + \mu'') + \frac{1}{2}\varepsilon^{-1}|v'' - \mu''|^2.$$

Fix an $r > 0$. Then for all $\varepsilon > 0$ small enough, we have

$$\underline{I}_\varepsilon(r) = \min_{v \in r\mathbb{S}^{d-1}} I_\varepsilon(v) \geq \min_{\substack{v \in r\mathbb{S}^{d-1}: \\ |v'' - \mu''| \leq \varepsilon^{1/3}}} I(v' + \mu'') \geq \min_{\substack{v \in r\mathbb{S}^{d-1}: \\ |v'' - \mu''| \leq \varepsilon^{1/3}}} \underline{I}\left(\sqrt{|v'|^2 + |\mu''|^2}\right),$$

where in the first equality we used no v such that $|v'' - \mu''| > \varepsilon^{1/3}$ contributes to the first minimum since $I_\varepsilon(v) > \frac{1}{2}\varepsilon^{-1/3}$ for such v . Finally, since $|v'|^2 = r^2 - |v''|^2$ for $v \in r\mathbb{S}^{d-1}$,

$$\underline{I}_\varepsilon(r) \geq \min_{\substack{v \in r\mathbb{S}^{d-1}: \\ |v'' - \mu''| \leq \varepsilon^{1/3}}} \underline{I}\left(\sqrt{r^2 + |\mu''|^2 - |v''|^2}\right) \geq \min_{|\delta| \leq 2\varepsilon^{1/3}|\mu''| + \varepsilon^{2/3}} \underline{I}\left(\sqrt{r^2 + \delta}\right).$$

Then, since $\underline{I}(r) \geq \underline{I}_\varepsilon(r)$ and \underline{I} is continuous at r , we obtain that $\underline{I}_\varepsilon(r) \rightarrow \underline{I}(r)$ as $\varepsilon \rightarrow 0+$ for every $r > 0$, as required. This is also true for $r = 0$ since $I_\varepsilon(0) \rightarrow I(0)$. \square

4. Proofs of the main results

4.1. Basic facts on large deviations

- Let $(Z_n)_{n \geq 1}$ be random elements of a Polish space \mathcal{M} equipped with a metric d , and let $\mathcal{J} : \mathcal{M} \rightarrow [0, \infty]$ be a lower semi-continuous function. We say that \mathcal{J} is *tight* if its sub-level sets $\{x \in \mathcal{M} : \mathcal{J}(x) \leq \alpha\}_{\alpha \geq 0}$ are compact. We say that the collection $(Z_n)_{n \geq 1}$ satisfies a *large deviations principle (LDP)*, in short) in \mathcal{M} with speed n and the *rate function* \mathcal{J} if for every Borel set $B \subset \mathcal{M}$,

$$(4.1) \quad \begin{aligned} - \inf_{x \in \text{int } B} \mathcal{J}(x) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in B) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in B) \leq - \inf_{x \in \text{cl } B} \mathcal{J}(x), \end{aligned}$$

where, as usual, we agree that $\inf_\emptyset = +\infty$. We assume throughout that \mathcal{J} is tight; so are all the rate functions considered in this paper. A Borel set $B \subset \mathcal{M}$ is called

regular for the rate function \mathcal{J} if the infima in (4.1) are equal. Since \mathcal{J} is tight, the infimum on the r.h.s. of (4.1) is always attained at some x .

- Assume that $B \subset \mathcal{M}$ is a closed set such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Z_n \in B) = - \inf_{x \in B} \mathcal{J}(x)$$

(e.g., we can take any regular closed set) and $B \cap \mathcal{D}_{\mathcal{J}} \neq \emptyset$. Then for any $\varepsilon > 0$,

$$(4.2) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(d(Z_n, x) \leq \varepsilon \text{ for some } x \in B\right) \\ \text{such that } \mathcal{J}(x) = \min_{y \in B} \mathcal{J}(y) \Big| Z_n \in B = 1.$$

This means that given the large deviations event $\{Z_n \in B\}$, the random elements Z_n asymptotically concentrate around the compact set of minimizers of the rate function \mathcal{J} over B . This follows from (4.1) since the conditioned event in (4.2) is $\{d(Z_n, \operatorname{argmin}_{x \in B} \mathcal{J}(x)) \leq \varepsilon\}$ and we have

$$\min_{x \in B} \mathcal{J}(x) < \inf_{x \in B: d(x, \operatorname{argmin}_{y \in B} \mathcal{J}(y)) \geq \varepsilon} \mathcal{J}(x).$$

The last inequality holds true since by tightness of \mathcal{J} , the infimum on the r.h.s. is attained on some $x \notin \operatorname{argmin}_{y \in B} \mathcal{J}(y)$ unless the minimum is taken over the empty set, in which case the r.h.s. is $+\infty$ and the inequality is still true.

- Denote by $C_0[0, 1] = C_0([0, 1]; \mathbb{R}^2)$ the space of continuous functions $h : [0, 1] \rightarrow \mathbb{R}^2$, i.e. planar curves, that satisfy $h(0) = 0$. We equip this space with the usual metric of uniform convergence. Denote by $AC_0[0, 1]$ its subspace of functions with absolutely continuous coordinates. Let $S_n(\cdot) \in C_0[0, 1]$ be the random piecewise linear functions that satisfy $S_n(k/n) := S_k$, $0 \leq k \leq n$, where $S_0 := 0$, and their values at the other points of $[0, 1]$ are defined by linear interpolation. Define the function $I_C : C_0[0, 1] \rightarrow [0, \infty]$ to be

$$(4.3) \quad I_C(h) := \begin{cases} \int_0^1 I(h'(t)) dt, & \text{if } h \in AC_0[0, 1]; \\ +\infty, & \text{otherwise.} \end{cases}$$

The following result, although stated in a different form, is due to Mogulskii [Mog76, Theorem 2, Part II]; there were earlier works in this direction by A.A. Borovkov. The exact statement presented here appears in the proof of Theorem 5.1.2 in book by Dembo and Zeitouni [DZ10].

THEOREM (Mogulskii's LDP). — Assume that X_1 is a random vector in \mathbb{R}^d , $d \geq 1$, such that $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^d$. Then the sequence of random functions $(S_n(\cdot)/n)_{n \geq 1}$ satisfies the LDP in $C_0[0, 1]$ with speed n and the tight rate function I_C .

- The above LDP for the trajectories $S_n(\cdot)/n$ readily implies that the random vectors $(S_n/n)_{n \geq 1}$ satisfy the LDP in \mathbb{R}^2 with speed n and the tight rate function $I_1(v) := \inf_{h: h(1)=v} I_C(h)$ for $v \in \mathbb{R}^2$. This follows by applying the *contraction principle* ([DZ10, Theorem 4.2.1]) and continuity of the mapping $h \mapsto h(1)$. Then $I_1 = I$ by Jensen's inequality:

$$(4.4) \quad I_C(h) = \int_0^1 I(h'(t)) dt \geq I(h(1)), \quad h \in AC_0[0, 1].$$

In particular, for any Borel set $B \subset \mathbb{R}^2$ that is regular for the rate function I , we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n/n \in B) = - \inf_{v \in B} I(v).$$

Since the rate function I is strictly convex, by (4.2) this implies that the trajectories $S_n(\cdot)$ that result in the large deviations event $\{S_n/n \in B\}$ are asymptotically linear, as in (2.7).

4.2. Main proofs

In this section we prove our main results Theorems 2.3, 2.11, and 2.13. The proofs follow the same idea of using classical geometric inequalities to solve the variational problems (2.4) and (2.9) and thus find the rate functions \mathcal{J}_P and \mathcal{J}_A .

Proof of Theorem 2.3.

(1) With a slight abuse of notation, denote by $P(h) := P(\text{conv}(\text{im } h))$ the perimeter of the convex hull of the image of a curve $h \in C_0[0, 1]$. This is a continuous functional on $C_0[0, 1]$, as follows from Cauchy’s formula (A.2). By

$$(4.5) \quad \text{conv} \left(\{S_n(t)\}_{0 \leq t \leq 1} \right) = \text{conv}(S_0, S_1, \dots, S_n),$$

one has

$$\frac{1}{2} P(S_n(\cdot)/n) = P_n/(2n).$$

This equality, Mogulskii’s LDP for trajectories of random walks (see Section 4.1), and the contraction principle ([DZ10, Theorem 4.2.1]) for continuous mappings yield that the sequence $(P_n/(2n))_{n \geq 1}$ satisfies an LDP in \mathbb{R} with speed n and the tight rate function

$$(4.6) \quad \mathcal{J}_P(x) := \inf_{h \in C_0[0, 1]: P(h) = 2x} I_C(h) = \min_{h \in AC_0[0, 1]: P(h) = 2x} I_C(h).$$

where, recall, I_C is given by (4.3). This implies (2.4). We used that the lower semi-continuous non-negative function I_C on $C_0[0, 1]$ has compact sub-level sets and therefore it always attains its infimum over the closed set $\{P(h) = 2x\}$.

The function \mathcal{J}_P is lower semi-continuous on \mathbb{R} as a rate function. It clearly satisfies $\mathcal{J}_P \leq \underline{I}$. To show that it is strictly increasing on $[|\mu|, r_{max}]$, take any $x \neq |\mu|$ from this interval and choose an $h \in AC_0[0, 1]$ such that $P(h) = 2x$ and $\mathcal{J}_P(x) = I_C(h)$. If $h'(t) = h(1)$ a.e. t , then $\mathcal{J}_P(y) \leq \underline{I}(y) < \underline{I}(x) = \mathcal{J}_P(x)$ for any $y \in [|\mu|, x)$, as required. Otherwise, for any $\varepsilon \in (0, 1)$,

$$\begin{aligned} & I_C((1 - \varepsilon)h + \varepsilon h(1) \cdot) \\ &= \int_0^1 I((1 - \varepsilon)h'(t) + \varepsilon h(1)) dt < \int_0^1 [(1 - \varepsilon)I(h'(t)) + \varepsilon I(h(1))] dt, \end{aligned}$$

by strict convexity of I on \mathcal{D}_I . Hence $I_C((1 - \varepsilon)h + \varepsilon h(1) \cdot) < I_C(h) = \mathcal{J}_P(x)$ by Jensen’s inequality (4.4). On the other hand, we have $P((1 - \varepsilon)h + \varepsilon h(1) \cdot) < P(h) = 2x$, which follows from Cauchy’s formula (A.2) and the relation

$$\text{conv} \left(\{(1 - \varepsilon)h(t) + \varepsilon h(1)t\}_{0 \leq t \leq 1} \right) \subsetneq \text{conv}(\{h(t)\}_{0 \leq t \leq 1}).$$

The strict inequalities above imply strict monotonicity of \mathcal{J}_P on $[|\mu|, r_{max}]$.

(2) The main task is to find the minimum in (4.6) and its minimizers.

First consider the case $[0, |\mu|]$.

For any function $h \in AC_0[0, 1]$, it clearly holds $P(h) \geq 2|h(1)|$. Then by Jensen’s inequality (4.4) and the fact that \underline{I} is decreasing on $[0, |\mu|]$, for any $x \in [0, |\mu|]$.

$$\begin{aligned} \mathcal{J}_P(x) &= \min_{h: P(h)=2x} I_C(h) \geq \min_{h: P(h)=2x} I(h(1)) \geq \min_{h: |h(1)| \leq x} I(h(1)) \\ &= \min_{r \leq x} \underline{I}(r) = \underline{I}(x). \end{aligned}$$

These inequalities are actually equalities since

$$(4.7) \quad I_C(x\ell t) = \underline{I}(x), \quad P(x\ell t) = 2x, \quad x > 0, \ell \in \underline{\Lambda}_x.$$

Hence $\mathcal{J}_P = \underline{I}$ on $[0, |\mu|]$. Moreover, recalling that $H_P(x) = \{h : I_C(h) = \underline{I}(x), P(h) = 2x\}$,

$$(4.8) \quad H_P(x) = \{t \mapsto x\ell_x t\}, \quad x \in (r_{min}, |\mu|].$$

Indeed, the facts that \underline{I} is strictly decreasing on $(r_{min}, |\mu|]$ and that Jensen’s inequality (4.4) for the strictly convex rate function I turns into equality only on functions with a.e. constant derivative, imply that the minimum in (4.6) is attained only on functions $h \in AC_0[0, 1]$ that satisfy $P(h) = 2|h(1)| = 2x$, that is $h(t) = x\ell t$ for some $\ell \in \mathbb{S}^1$. The unique function h of this form that satisfies the equality $I_C(h) = \underline{I}(x)$ corresponds to the direction ℓ_x .

The equality in (2.5) for $x \in (r_{min}, |\mu|]$ now follows from the LDP for the perimeters $(P_n/(2n))_{n \geq 1}$ proved in Part (1). In fact, we have $\mathcal{J}_P = \underline{I}$ on $(r_{min}, |\mu|]$. On this interval \underline{I} is decreasing and convex (see Lemma 2.1(b)), hence continuous, and so the set $[0, x]$, which corresponds to the event $\{P_n \leq 2xn\}$, is regular for the rate function \mathcal{J}_P .

The claim in (2.5) for $x = r_{min}$ holds trivially by $P_n \geq 2r_{min}n$ a.s.

Consider now the case $[|\mu|, \infty)$.

Our main estimate follows from the inequality $I(v) \geq \text{conv} \underline{I}(|v|)$, $v \in \mathbb{R}^2$, and Jensen’s inequality applied with the convex function $\text{conv} \underline{I}$. For any $h \in AC_0[0, 1]$, we have

$$(4.9) \quad \begin{aligned} I_C(h) &= \int_0^1 I(h'(t))dt \geq \int_0^1 \text{conv} \underline{I}(|h'(t)|)dt \geq \text{conv} \underline{I} \left(\int_0^1 |h'(t)|dt \right) \\ &= \text{conv} \underline{I}(\text{Var}(h)), \end{aligned}$$

where $\text{Var}(h)$ denotes the total variation, i.e. the length, of a curve $h \in C_0[0, 1]$.

Now use the following well-known inequality (see Corollary A.2 in the Appendix), which is even referred to as geometric “folklore”: $\text{Var}(h) \geq \frac{1}{2}P(h)$ for any $h \in C_0[0, 1]$ of bounded variation. Since the function \underline{I} increases on $[|\mu|, r_{max}]$, so does its largest

convex minorant $\text{conv } \underline{I}$. With the above, from (4.9) we get: for $x \geq |\mu|$,

$$(4.10) \quad \mathcal{J}_P(x) = \min_{h: P(h)=2x} I_C(h) \geq \min_{h: \text{Var}(h) \geq x} I_C(h) \\ \geq \min_{h: \text{Var}(h) \geq x} (\text{conv } \underline{I}(\text{Var}(h))) \geq \min_{r \geq x} (\text{conv } \underline{I}(r)) = \text{conv } \underline{I}(x).$$

Using (4.7) for an upper bound, this gives $\text{conv } \underline{I} \leq \mathcal{J}_P \leq \underline{I}$ on $[|\mu|, \infty)$.

We claim that if $\underline{I}(x) = \text{conv } \underline{I}(x)$ for $x \geq |\mu|$, then

$$(4.11) \quad H_P(x) = \{t \mapsto x\ell t, \ell \in \underline{\Lambda}_x\}.$$

We first note that by $\text{conv } \underline{I} \leq \mathcal{J}_P \leq \underline{I}$ and the assumption $\underline{I}(x) = \text{conv } \underline{I}(x)$, all inequalities in (4.10) are equalities. Then, since $\text{conv } \underline{I}$ is strictly increasing on $[|\mu|, r_{max})$, the infima in (4.10) are attained on the functions $h \in C_0[0, 1]$ that satisfy $\text{Var}(h) = \frac{1}{2}P(h) = x$. By Corollary A.2 in the Appendix, such functions have the form $h(t) = |h(t)|\ell$ a.e. t for some $\ell \in \mathbb{S}^1$ and satisfy $\text{Var}(h) = x$. Further, the second inequality in (4.9) is an equality iff $|h'(t)| \in [x_1, x_2]$ a.e. t , where $[x_1, x_2]$ is the maximal by inclusion interval that contains x and is such that the restriction of $\text{conv } \underline{I}$ on $[x_1, x_2]$ is affine. Finally, the first inequality in (4.9) is an equality for a function $h \in AC_0[0, 1]$ that satisfies the conditions above iff

$$|h'(t)| \in \{y \in [x_1, x_2] : I(y\ell) = \text{conv } \underline{I}(y)\} =: L_x \text{ a.e. } t$$

with the direction ℓ which was already fixed above. Since the rate function I is strictly convex, so is $I(\cdot\ell)$, hence $L_x = \{x\}$. Thus we obtained that $|h'(t)| = x$ a.e. t and by $I(x\ell) = \underline{I}(x)$, we have $\ell \in \underline{\Lambda}_x$. This finishes the proof of (4.11).

It remains to prove (2.6). In general, for an $x \in [|\mu|, r_{max}]$ we can not assure regularity of the set $[x, \infty)$ (corresponding to the event $\{P_n \geq 2xn\}$) for the rate function \mathcal{J}_P . The upper bound in (2.6) immediately follows from the LDP for the perimeters $(P_n/(2n))_{n \geq 1}$ we proved in Part 1 and the inequality $\text{conv } \underline{I} \leq \mathcal{J}_P$ (cf. the upper bound in (4.1) and (4.10), respectively). For the lower bound in (2.6), we consider two cases. If \underline{I} is continuous at x , then we use the inequality $\mathcal{J}_P \leq \underline{I}$ and the LDP for the perimeters (cf. the lower bound in (4.1)). If \underline{I} is discontinuous at x , then by Lemma 2.1 (c), the distribution of X_1 has atoms at the points of $x\underline{\Lambda}_x$, which must have equal weights satisfying $\underline{I}(x) = -\log \mathbb{P}(X_1 = x\ell)$ for $\ell \in \underline{\Lambda}_x$. Then

$$\mathbb{P}(P_n \geq 2xn) \geq \mathbb{P}(S_k = kx\ell, k = 1, \dots, n \text{ for some } \ell \in \underline{\Lambda}_x) = \#(\underline{\Lambda}_x) e^{-n\underline{I}(x)},$$

which gives the lower bound in (2.6). The proof of (2.6) is now finished.

(3) The claims follow from the general statement (4.2) combined with (4.8), (4.11) and using that $\max_{0 \leq k \leq n} |S_k/n - h(k/n)| \leq \max_{0 \leq t \leq 1} |S_n(t) - h(t)|$ for any $h \in C_0[0, 1]$. □

Proof of Theorem 2.11. Our argument is fully based on the ideas we developed in the proof of Theorem 2.3.

(1) Denote by $A(h)$ the area of the convex hull of a curve $h \in C_0[0, 1]$, i.e. $A(h) := A(\text{conv}(\text{im } h))$. It follows from the Steiner formula (A.3) that A is a continuous functional on $C_0[0, 1]$. From (4.5), one has

$$A(S_n(\cdot)/n) = A_n/n^2.$$

This equality, Mogulskii’s LDP for trajectories of random walks (see Section 4.1), and the contraction principle ([DZ10, Theorem 4.2.1]) for continuous mappings yield that the sequence $(A_n/n^2)_{n \geq 1}$ satisfies an LDP in \mathbb{R} with speed n and the tight rate function

$$(4.12) \quad \mathcal{J}_A(a) = \inf_{h \in C_0[0,1]: A(h)=a} I_C(h) = \min_{h \in AC_0[0,1]: A(h)=a} I_C(h),$$

where, recall, I_C is given by (4.3). This implies (2.9). We used that the lower semi-continuous non-negative function I_C on $C_0[0,1]$ has compact sub-level sets and therefore it always attains its infimum over the closed set $\{A(h) = a\}$.

Clearly, $\mathcal{J}_A(a) = 0$. Let us check that \mathcal{J}_A is strictly increasing on the set $\mathcal{D}_{\mathcal{J}_A}$. This assertion is trivial if this set is $\{0\}$, otherwise pick a positive $a \in \mathcal{D}_{\mathcal{J}_A}$. Then $\mathcal{J}_A(a) = I_C(h)$ for some $h \in C_0[0,1]$ such that $A(h) = a$. Clearly, h' is not constant a.e. on $[0,1]$ since otherwise $A(h) = 0$. Consider the function h_s such that $h_s(t) = (t/s)h(s)$ for $t \in [0,s]$ and $h_s = h$ on $[s,1]$, where $s \in (0,1]$; put $h_0 := h$. The area $A(h_s)$ decreases in s and satisfies $A(h_0) = A(h)$, $A(h_1) = 0$. By Jensen’s inequality, we also have $I_C(h_s) \leq I_C(h)$. From strict convexity of I , this inequality is strict if $A(h_s) < A(h)$. Since $A(h_s)$ is continuous in $s \in [0,1]$, these inequalities imply that $\mathcal{J}_A(a_1) < A(h) = \mathcal{J}_A(a)$ for any $a_1 \in [0,a)$.

Thus, the rate function \mathcal{J}_A is left-continuous on $\mathcal{D}_{\mathcal{J}_A}$ since it is lower semi-continuous and increasing. Then (2.10) follows from (and is easily seen to be equivalent to) the LDP for the areas $(A_n/n^2)_{n \geq 1}$. Finally, (2.11) holds by the general result (4.2).

(2) The isoperimetric inequality for convex hulls,

$$(4.13) \quad A(h) \leq \text{Var}(h)^2/(2\pi),$$

is valid for any function $h \in C_0[0,1]$ of bounded variation. This is Ulam’s version of the classical Dido problem, solved by Moran [Mor46]. We have $\underline{I} = \text{conv } \underline{I}$ by convexity of \underline{I} , which follows from rotational invariance of the distribution of X_1 . Then by (4.9) and (4.12), for $a \geq 0$

$$\mathcal{J}_A(a) = \min_{h \in C_0[0,1]: A(h)=a} I_C(h) \geq \min_{h: \text{Var}(h)^2 \geq 2\pi a} I_C(h) \geq \min_{h: \text{Var}(h) \geq \sqrt{2\pi a}} \underline{I}(\text{Var}(h)),$$

hence $\mathcal{J}_A(a) \geq \underline{I}(\sqrt{2\pi a})$ by (4.13). These three inequalities actually are equalities, with the minima attained only on the functions that parametrize half circles with the constant speed $\sqrt{2\pi a}$, and thus (2.12) holds true. In fact, the value of I_C on such a function h is exactly $\underline{I}(\sqrt{2\pi a})$. Since $\underline{I}(\sqrt{2\pi a})$ strictly increases for $a \in [0, r_{max}]$, it must be that $a = A(h) = \text{Var}(h)^2/(2\pi)$ and $|h'(t)|$ is constant for a.e. t , ensuring that the second inequality in (4.9) is an equality. And the isoperimetric inequality (4.13) is an equality only on parametrizations of semi-circles; see Tilli [Til10].

Finally, by Lemma 2.1 (b), the rate function $\mathcal{J}_A(a)$ is continuous on $[0, \infty)$, hence (2.10) is valid for every $a \geq 0$. □

Proof of Theorem 2.13. We need to find the rate function \mathcal{J}_A given by (2.9). For any $h \in AC_0[0, 1]$, by Jensen's inequality we have

$$\begin{aligned} I_C(h) &= \frac{1}{2} \int_0^1 |h'(t) - \mu|^2 dt \\ &= \frac{1}{2} \int_0^1 |h'(t)|^2 dt - h(1) \cdot \mu + \frac{1}{2} |\mu|^2 \geq \frac{1}{2} \text{Var}(h)^2 - h(1) \cdot \mu + \frac{1}{2} |\mu|^2, \end{aligned}$$

where the inequality is an equality iff $|h'(t)| = \text{Var}(h)$ for a.e. t . Hence, using that $\text{Var}(h)$ is invariant under rotations of the image of h about 0, we get

$$\begin{aligned} \mathcal{J}_A(a) &= \min_{h \in C_0[0, 1]: A(h)=a} I_C(h) \\ (4.14) \quad &= \min_{r \geq 0} \left(-r|\mu| + \min_{\substack{h: A(h)=a, \\ h(1)=r\mu/|\mu|}} \frac{1}{2} \text{Var}(h)^2 \right) + \frac{1}{2} |\mu|^2, \quad a \geq 0. \end{aligned}$$

Assume $a > 0$. It follows immediately from an approximation argument and the result by Pach [Pac78] for polygonal lines (see his Theorem 2 and the Remark just after it) that the above minimum over h with the fixed endpoint $h(1)$ is attained only on parametrizations h of circular arcs with $A(h) = a$. Denote by R the radius of such an arc and by 2φ its angle, where $R > 0$ and $0 \leq \varphi \leq \pi$. Then $\text{Var}(h) = 2\varphi R$, $\sin \varphi = r/(2R)$, and $A(h) = \varphi R^2 - \frac{1}{2}rR \cos \varphi$ in both cases $0 \leq \varphi \leq \pi/2$ and $\pi/2 \leq \varphi \leq \pi$. Due to the fact that $\varphi/\sin \varphi$ is strictly increasing on $[0, \pi]$, the mapping $(\varphi, R) \mapsto (r, V)$ is a bijection between the sets $[0, \pi] \times (0, \infty)$ and $\{(r, V) \in [0, \infty) \times (0, \infty) : r \leq V\}$. Hence (4.14) reduces to

$$(4.15) \quad \mathcal{J}_A(a) = \frac{1}{2} |\mu|^2 + 2 \min_{\substack{0 \leq \varphi \leq \pi, R \geq 0: \\ R^2 (\varphi - \frac{1}{2} \sin 2\varphi) = a}} (\varphi^2 R^2 - |\mu| R \sin \varphi).$$

Note that $R(\varphi) = \sqrt{a/(\varphi - \frac{1}{2} \sin 2\varphi)}$ satisfies $\frac{\partial R}{\partial \varphi} = -a^{-1} R^3 \sin^2 \varphi$. The values of the function $\varphi^2 R(\varphi)^2 - |\mu| R(\varphi) \sin \varphi$ at $0, \pi/2, \pi$ are respectively $+\infty, \pi a/2, \pi a$, hence this function attains its minimum at a critical point inside $(0, \pi)$ satisfying

$$2\varphi R^2 + 2\varphi^2 R \frac{\partial R}{\partial \varphi} = |\mu| \frac{\partial R}{\partial \varphi} \sin \varphi + |\mu| R \cos \varphi.$$

Dividing by R^4 and substituting the expression for $\frac{\partial R}{\partial \varphi}$ gives

$$\frac{2\varphi}{a} (\varphi - \sin \varphi \cos \varphi - \varphi \sin^2 \varphi) = \frac{|\mu|}{aR} (-\sin^3 \varphi + (\varphi - \sin \varphi \cos \varphi) \cos \varphi).$$

Then $2\varphi(\varphi \cos^2 \varphi - \sin \varphi \cos \varphi) = \frac{|\mu|}{R} (-\sin \varphi + \varphi \cos \varphi)$, and using that $\varphi \neq \tan \varphi$ on $(0, \pi/2)$,

$$(4.16) \quad R = \frac{|\mu|}{2\varphi \cos \varphi},$$

which is possible only when $\varphi \in (0, \pi/2)$. This gives

$$(4.17) \quad \frac{a}{|\mu|^2} = \frac{2\varphi - \sin 2\varphi}{8\varphi^2 \cos^2 \varphi}.$$

It easy to check that this equation has only one solution $\varphi \in [0, \pi/2)$ for every $a \geq 0$. In fact, the right-hand side of (4.17) equals zero at $\varphi = 0$ and $+\infty$ at $\varphi = \pi/2$, and its derivative

$$\frac{1}{2\varphi^2 \cos^3 \varphi} (\cos^2 \varphi \sin \varphi + \varphi^2 \sin \varphi - \varphi \cos \varphi)$$

is positive on $(0, \pi/2)$ by

$$\begin{aligned} \cos^2 \varphi \sin \varphi + \varphi^2 \sin \varphi - \varphi \cos \varphi &> \cos^2 \varphi \sin \varphi + \sin^3 \varphi - \varphi \cos \varphi \\ &= \sin \varphi - \varphi \cos \varphi = \cos \varphi (\tan \varphi - \varphi) > 0. \end{aligned}$$

Substituting (4.16) into (4.15) and using (4.17), we obtain

$$\mathcal{J}_A(a) = \inf_{h \in C_0[0, 1]: A(h)=a} I_C(h) = \frac{|\mu|^2}{2} \left(\frac{1}{\cos^2 \varphi} - \frac{2 \tan \varphi}{\varphi} + 1 \right) = 4\varphi a - \frac{1}{2} |\mu|^2 \tan^2 \varphi.$$

Clearly, this function is continuous in a , hence (2.10) is valid for every $a \geq 0$. The infimum is attained only at either of the two μ -axially symmetric curves

$$R(\sin(2\varphi t - \varphi) + \sin \varphi, \pm \cos(2\varphi t - \varphi) \mp \cos \varphi),$$

where φ and R are given by (4.16) and (4.17). This yields (2.14). □

4.3. The LDP's in continuous time

Here we obtain LDP's for convex hulls of Lévy processes by reduction to random walks.

Proof of Theorem 2.16. — First consider the perimeter P_T of the convex hull $C_T = \text{conv}(\{S_t\}_{0 \leq t \leq T})$ of the Lévy process $(S_t)_{t \geq 0}$. We shall compare it with the perimeter $P_{[T]}$ of the convex hull $C_{[T]} = \text{conv}(0, S_1, \dots, S_{[T]})$ of the random walk $(S_t)_{t \in \mathbb{N} \cup \{0\}}$.

It follows from Cauchy's formula (A.2) that

$$0 \leq (P_T - P_{[T]}) / (2\pi) \leq \max_{k \in \{0, 1, \dots, [T]\}} \sup_{k \leq t \leq k+1} |S_t - S_k| =: d_T,$$

where d_T is an upper bound for the Hausdorff distance between C_T and $C_{[T]}$. Let us estimate probabilities of large deviations of d_T/T . By stationarity of increments of $(S_t)_{t \geq 0}$, for every $\varepsilon > 0$ we have

$$\mathbb{P}(d_T \geq \varepsilon T) \leq ([T] + 1) \mathbb{P} \left(\sup_{0 \leq t \leq 1} |S_t| \geq \varepsilon T \right).$$

Put $\tilde{S}_t := S_t - t\mu$ for $t \geq 0$ (recall that $S_1 = X_1$) and let $\tilde{S}_t^{(1)}$ and $\tilde{S}_t^{(2)}$ be the coordinates of \tilde{S}_t is any orthonormal basis of \mathbb{R}^2 . Note that

$$\sup_{0 \leq t \leq 1} |S_t| \leq |\mu| + \sup_{0 \leq t \leq 1} |\tilde{S}_t| \leq |\mu| + \max_{i, j \in \{1, 2\}} \sup_{0 \leq t \leq 1} ((-1)^i \tilde{S}_t^{(j)}).$$

Denote $a := \varepsilon T - |\mu|$. Then for any $u > 0$ we get

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq t \leq 1} |S_t| \geq \varepsilon T\right) &\leq \sum_{i,j=1}^2 \mathbb{P}\left(\sup_{0 \leq t \leq 1} ((-1)^i \tilde{S}_t^{(j)}) \geq a\right) \\ &= \sum_{i,j=1}^2 \mathbb{P}\left(\sup_{0 \leq tv \leq 1} e^{u(-1)^i \tilde{S}_t^{(j)}} \geq e^{ua}\right). \end{aligned}$$

Since $(\tilde{S}_t)_{t \geq 0}$ is a zero-mean Lévy process in \mathbb{R}^2 , each of the four stochastic processes $((-1)^i \tilde{S}_t^{(j)})_{t \geq 0}$ is a right-continuous real-valued martingale. Then $(e^{u(-1)^i \tilde{S}_t^{(j)}})_{t \geq 0}$ are right-continuous positive sub-martingales, because $x \mapsto e^{ux}$ is a positive convex function of $x \in \mathbb{R}$. Hence, applying Doob’s maximal inequality (Revuz and Yor [RY99, Theorem 1.7, Chapter II]), we obtain

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq t \leq 1} |S_t| \geq \varepsilon T\right) &\leq \sum_{i,j=1}^2 e^{-ua} \mathbb{E} e^{u(-1)^i \tilde{S}_1^{(j)}} \\ &= \sum_{i,j=1}^2 \exp\left\{-\left(ua - \log \mathbb{E} e^{u(-1)^i \tilde{S}_1^{(j)}}\right)\right\}. \end{aligned}$$

Finally, if $a > 0$ (where $0 = \mathbb{E} \tilde{S}_1^{(j)}$), then optimizing the last expression over $u > 0$ yields

$$\mathbb{P}(d_T \geq \varepsilon T) \leq ([T] + 1) \sum_{i,j=1}^2 \exp\{-I_{i,j}(\varepsilon T - |\mu|)\},$$

where $I_{i,j}$ denotes the rate function of $(-1)^i \tilde{S}_1^{(j)}$.

Since the Laplace transform of S_1 is finite in \mathbb{R}^2 by the assumption, the Laplace transform of each of the random variables $(-1)^i \tilde{S}_1^{(j)}$ is finite in \mathbb{R} . This implies $\lim_{u \rightarrow \infty} I_{i,j}(u)/u = \infty$; see Rockafellar [Roc70, Theorems 8.5 and 13.3] or Vysotsky [Vys21a, Eqs. (5.4) and (5.5)]. Therefore, for every $\varepsilon > 0$, we have

$$(4.18) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \log \mathbb{P}(d_T/T > \varepsilon) = -\infty,$$

which means that the sequence of random variables $(d_T/T)_{T>0}$ is exponentially equivalent to 0 as $T \rightarrow \infty$ in the sense of [DZ10, Definition 4.2.10].

Finally, let us use that

$$\begin{aligned} \left|P_T/T - P_{[T]}/[T]\right| &\leq \left|P_T/T - P_{[T]}/T\right| + P_{[T]}(1/[T] - 1/T) \\ &\leq 2\pi d_T/T + (P_{[T]}/[T])/T, \end{aligned}$$

where the r.h.s. is exponentially equivalent to 0 as $T \rightarrow \infty$ by (4.18) and the fact that $(P_n/n)_{n \in \mathbb{N}}$ satisfies an LDP in \mathbb{R} with a tight rate function (by Theorem 2.3). Therefore, the sequences $P_T/(2T)$ and $P_{[T]}/(2[T])$ are exponentially equivalent as $T \rightarrow \infty$, hence they satisfy the same LDP by [DZ10, Theorem 4.2.13], as claimed.

As for the areas, the Steiner formula (A.3) yields

$$0 \leq (A_T - A_{[T]}) \leq P_{[T]}d_T + \pi d_T^2,$$

and it follows by the same argument as above that A_T/T^2 and $A_{[T]}/T^2$ are exponentially equivalent as $T \rightarrow \infty$ (use (4.18) and the facts that $(P_n/n)_{n \in \mathbb{N}}$ and $(A_n/n^2)_{n \in \mathbb{N}}$ satisfy LDPs with tight rate functions). Then it follows from Theorem 2.11 and [DZ10, Theorem 4.2.13] that $(A_T/T^2)_{T \geq 1}$ and $(A_{[T]}/[T]^2)_{T \geq 1}$ satisfy the same LDP, as claimed. \square

4.4. The LDP's under the Cramér moment assumption

Here we partially extend our main Theorems 2.3 and 2.11 under the weaker assumption $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$. We will use the contraction principle by Vysotsky [Vys21a].

Denote by $BV[0, 1] = BV([0, 1]; \mathbb{R}^2)$ the set of right-continuous functions of bounded variation from $[0, 1]$ to \mathbb{R}^2 . Denote by $A(h)$ and $P(h)$ respectively the area and the perimeter of $\text{conv}(h([0, 1]) \cup \{0\})$ of an $h \in BV[0, 1]$. This extends the definitions given in Section 4.2 for $h \in C_0[0, 1]$. Consider the functional

$$(4.19) \quad I_{BV}(h) := \sup_{\mathbf{t} \subset (0,1): \#\mathbf{t} < \infty} I_C(h^{\mathbf{t}}), \quad h \in BV[0, 1],$$

where $h^{\mathbf{t}}$ denotes the continuous function on $[0, 1]$ defined by linear interpolation between its values at $\mathbf{t} \cup \{0, 1\}$ that are given by $h^{\mathbf{t}}(s) := h(s)$ for $s \in \mathbf{t} \cup \{1\}$ and $h^{\mathbf{t}}(0) := 0$. This functional satisfies $I_{BV} = I_C$ on $AC_0[0, 1]$; see [Vys21a, Theorem 5.1], which gives an explicit and transparent formula for $I_{BV}(h)$ in terms of the Lebesgue decomposition of h .

PROPOSITION 4.1. — *Assume that X_1 is a random vector in the plane such that $0 \in \text{int } \mathcal{D}_{\mathcal{L}}$. Then the random variables $(P_n/(2n))_{n \geq 1}$ and $(A_n/n^2)_{n \geq 1}$ satisfy the LDP's in \mathbb{R} with speed n and the respective tight rate functions $\tilde{\mathcal{J}}_P$ and $\tilde{\mathcal{J}}_A$ given by*

$$(4.20) \quad \tilde{\mathcal{J}}_P(x) := \text{cl} \inf_{\substack{h \in BV[0,1]: \\ P(h) = 2x}} I_{BV}(h), \quad \tilde{\mathcal{J}}_A(x) := \text{cl} \inf_{\substack{h \in BV[0,1]: \\ A(h) = x}} I_{BV}(h), \quad x \geq 0.$$

These rate functions increase on $[|\mu|, \infty)$ and $[0, \infty)$, respectively. We always have $\tilde{\mathcal{J}}_P = \underline{I}$ on $[0, |\mu|]$. Moreover, $\tilde{\mathcal{J}}_P = \underline{I}$ if \underline{I} is convex. Also, we have $\tilde{\mathcal{J}}_A(a) = \underline{I}(\sqrt{2\pi x})$ for $x \geq 0$ if the distribution of X_1 is rotationally invariant.

Note that the monotonicity properties of $\tilde{\mathcal{J}}_P$ and $\tilde{\mathcal{J}}_A$ imply that the lower semi-continuous regularizations cl in (4.20) may change the values of the infima only at the discontinuity points.

Proof. — Let us equip $BV[0, 1]$ with the metric ρ equal the Hausdorff distance between the *completed graphs* of functions, defined by $\Gamma h := \{(t, x) : 0 \leq t \leq 1, x \in [h(t-), h(t)]\}$ for $h \in BV[0, 1]$, where $h(0-) := 0$. Note that Γh is a compact subset of $[0, 1] \times \mathbb{R}^2$ and it uniquely defines h , i.e. $\Gamma h_1 = \Gamma h_2$ for $h_1, h_2 \in BV[0, 1]$ implies $h_1 = h_2$. The total variation of an $h \in BV[0, 1]$, given by $\text{Var}(h) := \sup_{\mathbf{t} \subset (0,1): \#\mathbf{t} < \infty} \text{Var}(h^{\mathbf{t}})$, is simply the length of the spatial coordinate of any continuous bijective parametrization of Γh .

It follows from Steiner's and Cauchy's formulas (A.3) and (A.2) that the functionals A and P are continuous in the metric ρ and moreover, they are *uniformly* continuous

on the sets $\{h \in BV[0, 1] : \text{Var}(h) \leq R\}$ for every $R > 0$. Therefore, the contraction principle for the trajectories $S_n(\cdot)$ in $BV[0, 1]$, given by [Vys21a, Theorem 3.3] (which uses a metric longer than ρ , see [Vys21a, Eqs. (2.6) and (2.7)]), yields the LDPs stated with the respective rate functions given in (4.20).

The rest of the proof is identical to the ones of the corresponding parts of Theorems 2.3 and 2.11. We comment only on the differences. The monotonicity properties of $\tilde{\mathcal{J}}_P$ and $\tilde{\mathcal{J}}_A$ follow from equalities (4.19) and (4.20). We get only non-strict monotonicity since we are not claiming that the infima in (4.20) are always attained, as opposed to the main case $\mathcal{D}_{\mathcal{L}} = \mathbb{R}^2$.

Furthermore, by (4.19) and Jensen’s inequality, we have $I_{BV}(h) \geq I(h(1))$ for any $h \in BV[0, 1]$. Moreover, if \underline{I} is convex, we have $I_{BV}(h) \geq \underline{I}(\text{Var}(h))$. This follows from (4.9) and (4.19) using lower semi-continuity of \underline{I} (Lemma 2.1 (b)) if we choose an increasing sequence $(\mathbf{t}_n)_{n \geq 1}$ of finite subsets of $(0, 1]$ such that $I_C(h^{\mathbf{t}_n}) \rightarrow I_{BV}(h)$ and $\text{Var}(h^{\mathbf{t}_n}) \rightarrow \text{Var}(h)$ as $n \rightarrow \infty$. The two inequalities above for $I_{BV}(h)$ yield, as in the proof of Theorem 2.3, that $\inf_{h: P(h)=2x} I_{BV}(h) \geq \underline{I}(x)$ for any $x \in [0, |\mu|]$ and also for $x \geq |\mu|$ if \underline{I} is convex. Hence $\tilde{\mathcal{J}}_P(x) \geq \underline{I}(x)$ for such x since \underline{I} is lower semi-continuous. On the other hand, we have

$$\tilde{\mathcal{J}}_P(x) \leq \inf_{h \in BV[0, 1]: P(h)=2x} I_{BV}(h) \leq \inf_{\substack{h \in AC_0[0, 1]: \\ P(h)=2x}} I_{BV}(h) = \inf_{\substack{h \in AC_0[0, 1]: \\ P(h)=2x}} I_C(h) \leq \underline{I}(x),$$

where we used that $I_{BV} = I_C$ on $AC_0[0, 1]$. This yields the claims on $\tilde{\mathcal{J}}_P$.

Similarly, if \underline{I} is convex, which is surely the case when the distribution of X_1 is rotationally invariant, then we have $I_{BV}(h) \geq \underline{I}(\text{Var}(h))$ for $h \in BV[0, 1]$, hence $\inf_{h: A(h)=a} I_{BV}(h) \geq \underline{I}(\sqrt{2\pi a})$ for $a \geq 0$ by the same argument as in the proof of Theorem 2.11. Hence $\tilde{\mathcal{J}}_A(a) \geq \underline{I}(\sqrt{2\pi a})$ for $a \geq 0$. On the other hand, for rotationally invariant distributions of X_1 we have $\tilde{\mathcal{J}}_A(a) \leq \underline{I}(\sqrt{2\pi a})$, arguing as above for $\tilde{\mathcal{J}}_P(x) \leq \underline{I}(x)$. This yields the claim on $\tilde{\mathcal{J}}_A$. \square

Appendix A.

Perimeters. Throughout the paper, by the *perimeter* $P(C)$ of a non-empty convex set C on the plane we mean the length of its boundary unless C is a line segment, in which case $P(C)$ is its doubled length. Recall that a continuous curve in \mathbb{R}^d is *rectifiable* if it has finite length (equivalently, it has bounded variation).

The following simple proposition is proved in our separate note [AV17], which was initially motivated by the questions concerning the perimeter of the convex hulls considered in the present paper. For the reader’s convenience, we present the result here. Its main use here is in the corollary, which not only gives the “folklore” inequality for the half-perimeter but also specifies all instances when the equality is attained.

PROPOSITION A.1. — *Let γ be a rectifiable curve in \mathbb{R}^2 , and let Γ denote its convex hull. Then*

$$\text{length } \gamma \geq \text{per } \Gamma - \text{diam } \Gamma.$$

COROLLARY A.2. — *It holds that*

$$\text{length } \gamma \geq \frac{1}{2} \text{ per } \Gamma,$$

and equation can be attained only if γ parametrizes a line segment.

Remark A.3. — These statements remain valid if we replace \mathbb{R}^2 by \mathbb{R}^d (with any $d \geq 2$) and $\text{per } \Gamma$ by $\frac{dv_d}{v_{d-1}} W(\Gamma)$, where v_d denotes volume of a unit ball in \mathbb{R}^d and

$$(A.1) \quad W(\Gamma) := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} w_\ell(\Gamma) d\ell$$

is *mean width* of Γ , with $w_\ell(\Gamma)$ being width of Γ in the direction ℓ i.e. length of the projection of Γ on the line passing through the origin in the direction ℓ . The normalizing factor corresponds to mean width $\frac{2v_{d-1}}{dv_d}$ of a unit segment in \mathbb{R}^d .

It is easy to prove the remark using *Crofton's formula* (Schneider and Weil [SW08, Eq. (5.32)])

$$\text{length } \gamma = \frac{1}{v_{d-1}} \iint_{\mathbb{S}^{d-1} \mathbb{R}_+} n_\gamma(\ell, r) d\ell dr,$$

where $n_\gamma(\ell, r)$ denotes the number of intersections of γ with the hyperplane perpendicular to the direction ℓ at the distance r from the origin. Indeed, consider the closed curve γ' obtained by joining the end points of γ by a line segment. Almost every hyperplane intersecting Γ intersect γ' at least at two points since $\text{conv}(\gamma') = \Gamma$. It remains to use that $|\mathbb{S}^{d-1}| = dv_d$.

Note that Crofton's formula implies *Cauchy's formula* for the perimeter of the planar convex set Γ :

$$(A.2) \quad \text{per } \Gamma = \frac{1}{2} \int_{\mathbb{S}^1} w_\ell(\Gamma) d\ell.$$

Areas. Let $C \subset \mathbb{R}^2$ be a non-empty bounded convex set and let $B \subset \mathbb{R}^2$ be the closed unit ball centred at the origin. *Steiner's formula* ([SW08, Eq. (14.5)]) asserts that for every $r > 0$,

$$(A.3) \quad A(C + rB) = A(C) + P(C)r + \pi r^2,$$

where '+' stands for Minkowski addition of sets.

Measurability. Let us show that the perimeters and areas $(P_n)_{n \in \mathbb{N}}$, $(A_n)_{n \in \mathbb{N}}$, $(P_T)_{T > 0}$, $(A_T)_{T > 0}$ of the convex hulls, introduced in Sections 1 and 2.4, are measurable.

It follows from (A.2) and (A.3) that for every $n \in \mathbb{N}$, the mappings $(x_1, \dots, x_n) \mapsto P(\text{conv}(0, x_1, \dots, x_n))$ and $(x_1, \dots, x_n) \mapsto A(\text{conv}(0, x_1, \dots, x_n))$ are continuous from $\mathbb{R}^{2 \times n}$ to \mathbb{R} . Hence P_n and A_n are random variables.

Furthermore, for any $T > 0$ and a dense subset $\{t_k\}_{k \in \mathbb{N}}$ of $[0, T]$ that includes T ,

$$\begin{aligned} \text{cl } C_T &= \text{cl}(\text{conv}(\{S_t\}_{0 \leq t \leq T})) = \text{conv}(\text{cl}(\{S_t\}_{0 \leq t \leq T})) \\ &= \text{conv}(\text{cl}(\{S_{t_k}\}_{k \in \mathbb{N}})) = \text{cl}(\text{conv}(\{S_{t_k}\}_{k \in \mathbb{N}})) \quad \text{a.s.,} \end{aligned}$$

where the second and the fourth equalities hold true by [Roc70, Theorem 17.2], which applies because the trajectories of a Lévy process are bounded a.s. on any interval, and in the third equality we used that the trajectories are right-continuous and have left limits a.s. Then $\text{cl } C_T = \text{cl}(\cup_{k=1}^{\infty} \text{conv}(S_{t_1} \dots, S_{t_k}))$ by Carathéodory's theorem ([Roc70, Theorem 17.1]).

Hence, since the union on the r.h.s. is a convex set, we have

$$A_T = A(C_T) = A\left(\bigcup_{k=1}^{\infty} \text{conv}(S_{t_1} \dots, S_{t_k})\right) = \lim_{k \rightarrow \infty} A(\text{conv}(S_{t_1} \dots, S_{t_k})),$$

and by the above, A_T is measurable as a limit of measurable functions. Also, for any $\ell \in \mathbb{S}^1$,

$$w_{\ell}(C_T) = w_{\ell}\left(\bigcup_{k=1}^{\infty} \text{conv}(S_{t_1} \dots, S_{t_k})\right) = \lim_{k \rightarrow \infty} w_{\ell}(\text{conv}(S_{t_1} \dots, S_{t_k})),$$

which yields $P_T = \lim_{k \rightarrow \infty} P(\text{conv}(S_{t_1} \dots, S_{t_k}))$ by (A.2) and the monotone convergence theorem. Hence P_T is measurable as a limit of measurable functions.

Acknowledgements

We are grateful to Andrew Wade for bringing the perimeter problem to our attention, and to Endre Makai for referring us to the paper [Pac78] by János Pach. We wish to thank Fedor Petrov for showing us a simple proof of Proposition 3.2(a). We are indebted to the anonymous referees for their comments and the suggestion to include continuous time results.

BIBLIOGRAPHY

- [AKMV20] Gerold Alsmeyer, Zakhar Kabluchko, Alexander Marynych, and Vladislav Vysotsky, *How long is the convex minorant of a one-dimensional random walk?*, Electron. J. Probab. **25** (2020), article no. 105. ↑1165
- [AV17] Arseniy Akopyan and Vladislav Vysotsky, *On the lengths of curves passing through boundary points of a planar convex shape*, Am. Math. Mon. **124** (2017), no. 7, 588–596. ↑1197
- [Ber96] Jean Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, 1996. ↑1175
- [BM13] Aleksandr A. Borovkov and Anatolii A. Mogulskii, *Large deviation principles for random walk trajectories. II*, Theory Probab. Appl. **57** (2013), no. 1, 1–27. ↑1177
- [BN78] Ole Barndorff-Nielsen, *Information and exponential families in statistical theory*, John Wiley & Sons, 1978. ↑1167
- [BNB63] Ole Barndorff-Nielsen and Glen Baxter, *Combinatorial lemmas in higher dimensions*, Trans. Am. Math. Soc. **108** (1963), 313–325. ↑1165
- [CFG91] Hallard T. Croft, Kenneth J. Falconer, and Richard K. Guy, *Unsolved problems in geometry*, Problem Books in Mathematics, Springer, 1991. ↑1176
- [CHM15] Gunnar Claussen, Alexander K. Hartmann, and Satya N. Majumdar, *Convex hulls of random walks: large-deviation properties*, Phys. Rev. E **91** (2015), no. 5, article no. 052104. ↑1166

- [DZ10] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer, 2010, corrected reprint of the second (1998) edition. ↑1176, 1188, 1189, 1192, 1195, 1196
- [GSO16] Georg Glaeser, Hellmuth Stachel, and Boris Odehnal, *The Universe of Conics: From the ancient Greeks to 21st century developments*, Springer, 2016. ↑1185
- [Kho92] Davar Khoshnevisan, *Local asymptotic laws for the Brownian convex hull*, Probab. Theory Relat. Fields **93** (1992), no. 3, 377–392. ↑1166
- [KL98] James Kuelbs and Michel Ledoux, *On convex limit sets and Brownian motion*, J. Theor. Probab. **11** (1998), no. 2, 461–492. ↑1166
- [Mog76] Anatolii A. Mogulskii, *Large deviations for the trajectories of multidimensional random walks*, Theory Probab. Appl. **21** (1976), no. 2, 300–315. ↑1188
- [Mor46] Patrick A. P. Moran, *On a problem of S. Ulam*, J. Lond. Math. Soc. **21** (1946), 175–179. ↑1173, 1192
- [MW16] Ilya Molchanov and Florian Wespi, *Convex hulls of Lévy processes*, Electron. Commun. Probab. **21** (2016), article no. 69. ↑1166
- [MW18] James McRedmond and Andrew R. Wade, *The convex hull of a planar random walk: perimeter, diameter, and shape*, Electron. J. Probab. **23** (2018), article no. 131. ↑1165
- [Pac78] János Pach, *On an isoperimetric problem*, Stud. Sci. Math. Hung. **13** (1978), 43–45. ↑1174, 1193, 1199
- [Roc70] R. Tyrrell Rockafellar, *Convex analysis*, Princeton University Press, 1970. ↑1167, 1177, 1178, 1179, 1183, 1195, 1199
- [RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer, 1999. ↑1195
- [SS93] Timothy Law Snyder and J. Michael Steele, *Convex hulls of random walks*, Proc. Am. Math. Soc. **117** (1993), no. 4, 1165–1173. ↑1165, 1166
- [SW61] Frank Spitzer and Harold Widom, *The circumference of a convex polygon*, Proc. Am. Math. Soc. **12** (1961), 506–509. ↑1165
- [SW08] Rolf Schneider and Wolfgang Weil, *Stochastic and integral geometry*, Probability and Its Applications, Springer, 2008. ↑1176, 1198
- [Til10] Paolo Tilli, *Isoperimetric inequalities for convex hulls and related questions*, Trans. Am. Math. Soc. **362** (2010), no. 9, 4497–4509. ↑1176, 1192
- [Vys21a] Vladislav Vysotsky, *Contraction principle for trajectories of random walks and Cramér’s theorem for kernel-weighted sums*, ALEA, Lat. Am. J. Probab. Math. Stat. **18** (2021), 1103–1125. ↑1177, 1195, 1196, 1197
- [Vys21b] ———, *When is the rate function of a random vector strictly convex?*, Electron. Commun. Probab. **26** (2021), article no. 41. ↑1167
- [VZ18] Vladislav Vysotsky and Dmitry Zaporozhets, *Convex hulls of multidimensional random walks*, Trans. Am. Math. Soc. **370** (2018), no. 11, 7985–8012. ↑1165, 1176, 1177
- [WX15a] Andrew R. Wade and Chang Xu, *Convex hulls of planar random walks with drift*, Proc. Am. Math. Soc. **143** (2015), no. 11, 433–445. ↑1165
- [WX15b] ———, *Convex hulls of random walks and their scaling limits*, Stochastic Processes Appl. **125** (2015), no. 11, 4300–4320. ↑1165

Manuscript received on 15th December 2019,
revised on 30th September 2020,
accepted on 5th January 2021.

Recommended by Editor L. Chaumont.
Published under license CC BY 4.0.



This journal is a member of Centre Mersenne.



Arseniy AKOPYAN
Institute for Information
Transmission Problems RAS,
Bolshoy Karetny per. 19,
Moscow, 127994, (Russia)
akopjan@gmail.com

Vladislav VYSOTSKY
University of Sussex,
Pevensey 2 Building, Falmer Campus,
Brighton BN1 9QH, (United Kingdom)
and St. Petersburg Department
of Steklov Mathematical Institute,
Fontanka 27, 191011
St. Petersburg, (Russia)
v.vysotskiy@sussex.ac.uk